DOCUMENT RESUME

ED 297 428 EA 020 141

TITLE	Assessment of Educational Achievement in Asia and the Pacific. Report of a Seminar (Tokyo, Japan, January 20-February 1, 1986).
INSTITUTION	National Inst. for Educational Research, Tokyo (Japan).
PUB DATE	Aug 86
NOTE	239p.
PUB TYPE	Collected Works - Conference Proceedings (021)
EDRS PRICE	MF01/PC10 Plus Postage.
DESCRIPTORS	Comparative Analysis; *Developing Nations; *Educational Assessment; Educational Trends; Elementary Secondary Education; Equal Education; *Evaluation Criteria; Evaluation Utilization; *Foreign Countries; Research and Development; Research Needs; Student Development
IDENTIFIERS	Australia; Bangladesh; China; India; Japan; Malaysia; Nepal; New Zealand; Pakistan; Papua New Guinea; Philippines; Singapore; South Korea; Sri Lanka; Thailand

ABSTRACT

The High Level Seminar on Educational Achievement in Asia and the Pacific was organized in collaboration with the UNESCO Regional Office for Education in Asia and the Pacific (Bangkok, Thailand). The objectives were to exchange information and experiences on educational assessment at the primary and secondary school levels; to review major research and surveys on educational achievement and their findings conducted in the participating countries; and to identify common trends of research on educational achievement in the region. Part 1 of this report addresses in general the status, issues, and trends in the assessment of educational achievement in the Asia-Pacific region. After a preamble, the following topics are considered: (1) purposes of assessment; (2) structure and methods of assessment; (3) domains of assessment (cognitive, psychomotor, and affective); (4) equity and excellence; and (5) monitoring levels of assessment. A summary of recommendations and concluding remarks follow. Part 2 consists of reports from each of the participating countries: Australia, China, India, Indonesia, Japan, Malaysia, Nepal, New Zealand, Pakistan, Philippines, Republic of Korea, Sri Lanka, and Thailand. A list of participants is included. (TE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED297428

NIER

ASSESSMENT OF EDUCATIONAL ACHIEVEMENT IN ASIA AND THE PACIFIC

EA 020 141

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it

☐ Minor changes have been made to improve reproduction quality

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

T. Kanaya

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

National Institute for Educational Research, Tokyo, Japan

ASSESSMENT OF EDUCATIONAL ACHIEVEMENT IN ASIA AND THE PACIFIC

Report of a Seminar

Section for Educational Co-operation in Asia
National Institute for Educational Research (NIER)
Tokyo, Japan

August 1986

© NIER 1986

Published by

the Section for Educational Co-operation in Asia
National Institute for Educational Research (NIER)

5-22 Shimomeguro 6 Chome, Meguro-ku, Tokyo, Japan

CONTENTS

Introduction 1

Part I: Status, Issues and Trends in the Assessment of Educational Achievement in the Region

Preamble 3

1. Purposes of Assessment

1.1 Assessment in the Broader Context 4
1.2 Formative and Summative Assessment 5
1.3 Guidance 5
1.4 Student Certification 6
1.5 System Monitoring 7
1.6 Scholarships 7
1.7 Placement/Selection 8

2. Structure and Methods of Assessment

2.1 Responsibility for Assessment 8
2.2 Structure of Assessment 9
2.3 Tools and Techniques of Assessment 10
2.4 Criterion-/Norm-referenced Testing 10
2.5 Validity, Reliability and Usability of Tests . . . 11
2.6 Grading and Marking 13
2.7 Effects of Assessment 14
2.8 Teachers' Competencies in Assessment 16
Recommendation 17

3. Domains of Assessment

3.1 Cognitive Domain 17
3.2 Psychomotor Domain 18
3.3 Affective Domain 18
Recommendations 19

4. Equity and Excellence

4.1 Quality and Quantity 20
4.2 Succes and Retention/Retrieval 20
4.3 Disadvantaged Groups 21
Recommendations 22

5. Monitoring Levels of Assessment

5.1 Purpose of Monitoring 22
5.2 Types of Monitoring 22
5.3 Mastery and Minimum Learning 23
5.4 Uses of Research on Assessment 24
Recommendations 26

Summary of Recommendations 26

Concluding Remarks 28

Part II: Individual Country Reports

Australia (Dr. Sid F. Bourke) 31

China (Mr. Wan Yibing) 56

China (Mr. Yang Donglian) 64

India (Dr. P. N. Dave) 70

Indonesia (Dr. R. Ibrahim) 89

Japan (Mr. S. Kawanobe) 97

Malaysia (Mr. Mohd Rosli bin Mohd Ghazali) 109

Nepal (Mr. R. J. Thapa) 120

New Zealand (Mr. N. A. Reid) 139

Pakistan (Dr. Z. A. Ansari) 163

Philippines (Dr. M. M. Miguel) 172

Republic of Korea (Dr. Joung-Kyu Whang) 190

Republic of Korea (Mr. Lim Byung-Lang) 202

Sri Lanka (Mr. G. L. S. Nanayakkara) 205

Thailand (Mr. Winai Kasemsestha) 221

Annex: List of Participants 235

INTRODUCTION

Origin and Objectives

The High Level Seminar on Educational Achievement in Asia and the Pacific was organized by the National Institute for Educational Research (NIER) of Japan from 20 January to 1 February 1986 in collaboration with the Unesco Regional Office for Education in Asia and the Pacific in Bangkok. The seminar was one in a series launched in 1979 to promote educational research in the region of Asia and the Pacific. The beginning of this series, devoted to the consideration of educational research to the development of education by heads of educational research institutions in the region, was marked by a seminar on educational research in relation to educational reform held in 1979, followed by a seminar for the training of young educational researchers in 1981, a workshop on the development of research concerning the implementation and evaluation of reforms of educational content and methods in 1982, a seminar on educational research with special reference to the exchange of research outcomes in 1983 and a seminar on the contribution of interdisciplinary research to the development of education in 1984.

The objectives of the present seminar were :

1) to exchange information and experiences on the systems of the assessment of educational achievement at the primary and secondary school levels;

2) to review major researches and surveys on educational achievement and their findings conducted in the participating countries; and

3) to identify common trends of research on educational achievement in the region.

Participation and Organization

The seminar was attended by fifteen participants from thirteen countries in Asia and the Pacific. The countries represented were: Australia, China, India, Indonesia, Japan, Malaysia, Nepal, New Zealand, Pakistan, the Philippines, the Republic of Korea, Sri Lanka and Thailand. (for list of participants see Annex).

After the inaugural session, the seminar elected the following office bearers:

Chairperson:	Dr. P. N. Dave (India)
Vice-Chairperson:	Dr. M. M. Miguel (Philippines)
Rapporteurs:	Dr. S. F. Bourke (Australia)
	Mr. Mohd Rosli Mohd Ghazali (Malaysia)
	Mr. N. A. Reid (New Zealand)

The seminar met in plenary sessions, and after the presentation of individual reports (see Part II of this report), participants identified trends, issues and problems concerning assessment/evaluation of educational achievement.

The Report

The draft of Part I of this final report was presented and adopted at the closing session of the seminar and suggestions and modifications suggested at that session have been incorporated in this published version.

Part I

STATUS, ISSUES AND TRENDS IN THE ASSESSMENT OF EDUCATIONAL ACHIEVEMENT IN THE REGION

PREAMBLE

All the seminar participants contributed to Part I of this report which provides some theoretical rationale for educational assessment as well as representing the deliberations of the group. The involvement of all participants in preparing the report greatly assisted the flow of ideas at the seminar, but it meant that there are obvious differences in writing style between the sections. We hope that these differences are not too distracting to the reader.

Readers will also note that some topics are referred to more than once in Part I of the report. Such repetition results from the group approach adopted in preparing the report. While some redundancies were deleted from the draft report, the repetitions that remain have been retained because they address different aspects of the topic concerned and their deletion would result in the loss of important information.

Part I of the report is divided into five sections. As far as possible within each section, the structure used takes the form of describing the existing status of the aspect of assessment being discussed, identifying the current issues, describing similarities, differences and trends and, where possible, making recommendations. A summary of recommendations and some brief concluding remarks are also provided.

1. PURPOSES OF ASSESSMENT

1.1 Assessment in the Broader Context

In a very broad perspective, education may be conceived as an essential input for human resource development -- a citizen of the world transcending the national boundaries. Within a country, education has been regarded as preparation for an individual to become a fully functioning human being and a worthwhile citizen of his/her community/society.

Keeping in mind the national ethos, needs, concerns and problems, the curriculum is designed by each country to achieve these broad aims. Evaluation is viewed as an essential part of the curriculum which helps guide the learner towards the specific objectives, which have been derived from the broader national educational objectives.

Assessment is construed as an important element of the evaluation of the curriculum. Thus, it forms an imperative link with the school, society and the broader world.

The main purpose of assessment (often used interchangeably, but incorrectly, with the terms evaluation and measurement) is to collect evidence and information (data), to arrange them into a useful, interpretable form, and to make educational decisions. It is therefore not a uni-dimensional measurement, but a multi-trait or multi-method measurement.

If this purpose is in focus, three trends appear to be in vogue. Australia, Japan, New Zealand, the Republic of Korea and Thailand follow the multi-pronged approach of assessment while Malaysia and the Philippines use a good mixture of the uni- and multi-dimensional methods. The scope in China, India, Indonesia, Nepal, Pakistan and Sri Lanka seems to be more constricted. Even in these countries, the process of collecting information at the primary stage is much more broad-based and multi-dimensional than at the secondary stage.

Internal assessments are more comprehensive and sustained than those carried out externally. Up to the primary stage, usually only internal assessment is utilized for taking decisions regarding retention or promotion of the learner.

Although assessment precedes decision making, it cannot be totally isolated from it. Functionally two major linkages between education and the broader context have often been mentioned - entrance into the tertiary system and entering the world of work. Whether assessment is done internally or externally, all countries have mentioned the value of assessment for this linkage. However, India, Indonesia, Malaysia, Nepal, Pakistan, the Philippines, the Republic of Korea, Sri Lanka and Thailand reported that their assessment purposes have links with long-term national goals as well as with decisions about entrance into the tertiary system and the world of work which are based on school assessments.

If assessment is accepted as a multitrait-multidimensional process of gathering data, then it is felt that its scope needs to be enlarged in respect of all countries, particularly where assessment is too narrowly conceived.

1.2 Formative and Summative Assessment

Formative assessment is intended to diagnose the strengths and weaknesses of learners and to provide feedback to both the teacher and the learners so that the teaching-learning process becomes more efficient and productive. On the other hand, summative tests are used at the end of units, modules or programmes, or at specified times, e.g., quarterly, six-monthly and end-of-year. Associated closely with the intentions of these assessments, are the tools/procedures of assessment (discussed in Section 2.3). They are teacher-made or standardized tests and centralized or decentralized procedures.

All participating countries use both centralized and decentralized procedures. However, where internal assessment is emphasized as in Australia, Japan, New Zealand, the Philippines, the Republic of Korea and Thailand, the main purpose is diagnostic, and assessment is implemented by less formal means. In the remaining countries, where external assessment (examinations) is more important, the purpose is decision making or judging, using more formal means. The former is a decentralized procedure and the latter more centralized.

In the countries where external examinations dominate the curriculum, there is a dire need to make more use of formative assessment, while simultaneously decentralizing the process of assessment to the teacher and using a variety of informal techniques, such as observation, checklists and rating scales. However, it should be accomplished in such a way that the rhythm and balance of the existing assessment system is not disturbed unduly.

1.3 Guidance

As the main purpose of assessment is data gathering for decision making and diagnosis, guidance for taking corrective measures and making appropriate choices from among various alternatives of subjects and vocations acquires utmost importance. In the lower grades such information should help a teacher to remove impediments originating from socio-emotional problems.

The participating countries can, in general, be divided into two groups on the basis of centralization of the system of guidance. Countries which follow a decentralized system of internal assessment appear to gather and utilize information with an emphasis on guiding the development of the child. Countries which follow a centralized system need to be particularly careful to recognize individual needs and help to overcome personal, academic or emotional difficulties of individual students.

All participating countries in one form or another use information from internal or external assessments for selection of vocations/streams or screening students for professional courses at the secondary level. However, by and large, academic achievement appears to be the deciding factor for selection/screening of candidates, giving small or even no weighting to interests and aptitudes.

Guidance, particularly career guidance, has not been treated as sufficiently important in most countries. There is a need to have a more comprehensive profile assigning due weighting to aptitudes and interests, particularly in the countries where the system is highly competitive in the academic area.

1.4 Student Certification

A certificate should contain information in respect of the overall growth of the learner. Thus the objective of certification should be to communicate an accurate profile of the levels of achievement of the student to the teachers, parents, employers or the tertiary institutes, and most importantly to the learners themselves.

In all countries certificates are issued at the end of primary, or at the secondary levels, or both. In China , Indonesia, Japan, Nepal, the Republic of Korea, the Philippines and Thailand, the certificates are given at the both primary and secondary levels. The other countries issued their certificates at the secondary level only. For those countries issuing certificates at the primary level, external assessment is used. Certification at the secondary school level is fully internally assessed in Australia (except at Grade 12 in some states), Japan, the Philippines, the Republic of Korea, while in New Zealand certification is mixture of external and internal. For the other countries, certificates are issued on the basis of external assessments at the secondary level. Details of the grade levels at which assessments are made and certificates issued may be studied in the individual country reports (see Part II).

Certificates generally do not contain a holistic picture of the total achievement of the learners. In many countries where assessments are external, more often than not, marks/grades are shown for subjects, with all non-cognitive attainments excluded. However, it is considered desirable that certificates be issued at least at the end of the secondary stage of school education.

1.5 System Monitoring

In the operation of a curriculum, the school plays an extremely vital role. While the subject-based curriculum and associated assessments are carried out within a classroom, a totality of experiences is provided by the school. Therefore, there is as much need for monitoring the school as there is for monitoring the development of an individual. Put differently, in the context of national educational objectives, it is pertinent to ascertain the performance of a school within the system in the same way that the performance of an individual student within a school is assessed.

In this respect, no national institutionalized system of monitoring using formal assessments is found in Australia, China, India, Japan, New Zealand or Pakistan, although regional systems exist in some cases. Whereas Nepal and Sri Lanka reported the operation of such a system to some extent, Indonesia, Malaysia, the Philippines, the Republic of Korea and Thailand have a clear institutionalized mechanism for such monitoring. More detailed information on the nature of system monitoring is provided in Section 5 of this part of the report.

Many countries believe that, whether assessment is internal or external, decentralized or centralized, national monitoring is desirable in order that students in a widely scattered network of schools receive more or less uniform access to and opportunities in education. In the age of the computer, it should not be difficult to establish such a system of data storing and retrieval. However, in some countries where equal access to education has been largely achieved, the arguments in favour of a national monitoring system are considered to be less persuasive.

1.6 Scholarships

While recognizing the diagnostic purpose of assessment for remediating the weaknesses/difficulties of the learner, it is equally important that an educational system pays due attention to the development of talent in all walks of life.

Scholarships are one way of assisting talented students. The country reports indicate that as far as the identification and nurturing of talent in the cognitive area is concerned (except in Japan and Malaysia), a mechanism exists for providing facilities to this special group. Such a provision is sometimes made throughout schooling, but more often at the end of the secondary level of education. India is planning to create such an institution at the end of Class 5 on a national scale by establishing model schools in each district. However, in the non-cognitive area, no such concerted effort is discernible.

Irrespective of the country, there is a need to give proper stimulation for accelerating the level of attainment of the brilliant and accomplished students in the cognitive and non-cognitive areas. This may be done formally through scholarships or the rapid promotion of such students, or informally through the provision of extension work within special and mainstreamed classes.

1.7 Placement/Selection

These two linkages with the broader world relate to entry into the world of work and to tertiary education. Placement/selection of students may also be employed within the schooling system. The main purpose underlying placement is to guide the individual student to make a realistic and wise choice of educational streams and vocations or professional courses.

All participating countries utilize external assessment or guidance tests for helping students make choices or to place or admit them to one or other stream at the terminal stage of secondary education. Assessment is also used by the employer in selecting candidates for jobs. Thailand, in addition to placement, utilizes examination results for ability grouping in classrooms. Placement as early as the end of Grade 6 is done in China.

In general, assessment for placement/admission is found to be extremely competitive, where potential is judged mostly on academic or scholastic achievement. This appears to lack balance because affective characteristics play an important role in achieving success in a chosen field. A more balanced assessment needs to be developed for proper placement.

2. STRUCTURE AND METHODS OF ASSESSMENT

2.1 Responsibility for Assessment

In the participating countries, there are generally two loci of control over responsibility for assessment. The first type is where an individual school has total jurisdiction over its assessment programmes which include the whole gamut of test planning, construction, administration, data processing and dissemination of results. This type is referred to as internal assessment. The second type is external assessment where total responsibility for the assessment programme is located outside the school, and is in the hands of a government department, institution or examining board.

Internal assessment is employed by teachers in all countries with some variations in terms of the frequency with which the assessments are conducted. Most assessments consist of teacher-made tests and these are used mainly to provide feedback on the teaching-learning process. Assessments of this type are held periodically during the school year or at the end of each teaching module or unit. The results of such assessments are confined to uses within the school or as feedback to parents, often in the form of written school reports. However, in some countries at the senior secondary level (Australia,Indonesia and New Zealand), the results of a school-based internal assessment accounts for all or a proportion of the total assessment.

External assessment can be categorized into two kinds based on its extent, or how widely the assessment is administered. One kind is essentially small scale; they are localized programmes which may cover a cluster of schools, or schools within a district, division, or region. The second kind is a nation-wide administration, usually conducted by a central agency. Assessments are summative in nature and are administered at specific grade/class levels. For most countries, these are conducted in the secondary or middle level in the school system. There are some countries, however, which conduct nation-wide assessment programmes at the primary or elementary level as well.

2.2 Structure of Assessment

The main issue surfacing in internal assessment concerns the psychometric properties of the instruments used. These measures, which are usually teacher-made, are frequently suspect on several grounds in terms of quality. This is due almost entirely to teachers' lack of competency and/or experience in test development and allied knowledge, or to teachers having insufficient time to prepare the instruments properly.

External assessments, on the other hand, encounter fewer problems of this kind since most of the measures employed are systematically and expertly planned and developed. However, they may create other problems. The backwash effects of these assessments have created considerable anxiety and tension among students, and consequently have become a concern to educators and the general public alike.

Several methods have been introduced in order to reduce the problems encountered in both internal and external assessments. Currently, the trend in terms of control over assessment has been towards a shared responsibility between teachers, schools and external authorities. This change is being achieved through mixed internal and external assessment procedures or a decentralization where assessment of students' achievement have become increasingly school-based, currently more especially at the primary/elementary level.

The issue of whether to decentralize or centralize certification examinations exists in most countries. The practice of having a national agency, or Board or Ministerial testing committee to conduct selection and administration of examinations of secondary school graduates seeking entry to tertiary education is common to most.

2.3 Tools and Techniques of Assessment

In assessing educational achievement, most teachers use teacher-made measures which encompass objective tests, written and oral questions, performance tests, essays, short-answer tests, laboratory work, assignments, projects and the like. Informal and formal observations of individuals or groups, performance ratings and checklists and anecdotal records are also used for certain subjects in several of the participating countries.

An obvious omission, revealed in most country reports, was the lack of specific diagnostic testing for detecting individual learning difficulties. Where diagnostic measures were employed, they tended to be very broad.

Although there are wide variations in the tools and techniques employed, both in terms of range and popularity, uses for specific purposes are reported by some countries. In the Philippines, for example, attitude measures are used in civics and culture, social studies, work education and physical education. Psychological tests for guidance, remedial learning and research are used in the Republic of Korea and in other countries by appropriately qualified personnel.

Australia and New Zealand have published item banks, most often in science and mathematics. Pakistan has developed item banks for all major subjects in Classes 10 and 12. India, Indonesia, Nepal, the Philippines, the Republic of Korea and Sri Lanka are all currently developing question or item banks.

2.4 Criterion-/Norm-referenced Testing

Implicitly, most teacher-devised tests for classroom use are criterion, or content, or objective-referenced. Regardless of the part of the world in which they are working, all teachers are vitally concerned with what students are able to do presently (as a basis for future action); they wish to know how their students perform on the intellectual tasks incorporated in test questions which they set to assess learning. And whether or not they have formally stated criteria, competency cut-offs, or specific learning objectives and so on in planning and constructing their tests, teachers almost always will have an internalized notion of these requirements.

To a degree, however, such understandings of student behaviour and attainments are also norm-referenced. Teachers will have a relatively comprehensive and accurate idea of what students of a given age or class level are _typically_ capable of and are able to recognize what constitutes poor, and at the other end of the continuum, superior achievement.

Hence, the two types of assessment (C-R and N-R) need not be clearly delineated as is sometimes considered necessary. In most countries, classroom teachers would employ both kinds of assessments in their armoury of testing techniques, but perhaps unconsciously so, relying more on commonsense and intuition rather than a deep understanding.

Those relatively few teachers with a thorough knowledge of the attributes of N-R and C-R testing will sensibly employ whichever kind is appropriate for their assessment purposes and choose that which better achieves those purposes.

On the other hand, most large-scale examining and system surveying is, by its very nature, norm-referenced. Students are ranked on the basis of demonstrated achievement, regardless of the level of competency reached. The emphasis is not on what the student is able to do (although this may, of course, be determined by scrutinizing the content and student success on the questions making up the test), but instead on where he or she is ranked in comparison with age or class peers. Almost all participating countries that have systems employing external examinations for certification, awards, accreditation, matriculation or promotion, conduct tests that rank students in order of merit with cutting scores assigned at various points on the scale, depending on the purpose(s) of the examination(s).

Standardized achievement or scholastic aptitude tests are traditionally norm-referenced although they may present other types of information also, e.g., the New Zealand Progressive Achievement Tests provide content-referenced scores for some basic skills tests. Where they are readily available, tests of this kind are used by teachers in several countries (Australia, Japan, New Zealand and the Philippines), usually on a voluntary basis.

In those countries that have moved, or intend moving, towards a greater degree of school-based assessment with increased teacher autonomy in measurement matters, and/or are phasing out or abolishing external examinations, the trend is towards objective/domain-referenced and/or criterion-referenced school-based testing. Australia and New Zealand are clearly following this direction at the senior secondary school level. Other countries, such as India and the Philippines, where notions of 'mastery learning' and 'minimal competencies' have been incorporated in the system's prevailing educational philosophy are also of necessity moving away from traditional N-R assessment models to C-R ones.

2.5 Validity, Reliability and Usability of Tests

For all countries, at the classroom level, teacher-made tests are rarely checked for validity (relevance) or reliability (consistency), either at the developmental stage or after the test has been given. Hence, little is known regarding the important psychometric properties of these commonly used tests.

However, the prevailing opinion in all countries is that while teacher-made tests attempt to assess relevant outcomes of school instruction and are usually acceptable in terms of face validity, content validity (the most important type for achievement tests) frequently suffers for several reasons. These are unrepresentative sampling of what has been taught, inappropriate content emphases, restricted skill coverage, stereotyped questions measuring only those aspects of the curriculum most easily assessed, and an emphasis on recall and 'lower' cognitive processes to the exclusion of such abilities as evaluation, application of understanding and creativity. Thus, while content validity may reach satisfactory levels in some cases, it could undoubtedly be improved through increased teacher knowledge of assessment matters and with additional attention being paid to accepted techniques of test development. Where attempts have been made by classroom teachers and examiners to develop measures according to accepted psychometric principles for test construction, as, for example, in the senior levels of the school in the Philippines and Sri Lanka, estimates of validity and reliability have reached moderate to high levels.

Generally, for all participating countries, much more effort is devoted to matters of validity and reliability in the preparation of external examinations (where 'life chances' frequently rest on the results), and for measures employed in system monitoring. More knowledgeable and experienced teachers and/or people trained in educational assessment/psychometrics are typically involved in developing such measures. These are also usually checked, either judgementally or, less often, empirically, before they were administered to students. Time pressures encountered in test preparation (as in Indonesia and several other countries), or difficulties in ensuring efficient distribution, (e.g., Nepal), and occasionally breaches in test security with a need to prepare an equivalent test (as in India), etc., mean that the ideal levels of question/test scrutiny and excellence in test development are not reached, with concomitant deletarious effects on both validity and reliability.

The reliance on essay-type examinations in some countries, e.g., India and Pakistan, and the subjectivity inherent in this assessment technique and consequent lowered reliability, is an unresolved issue. While the essay is invaluable for measuring skills in synthesizing and in the presentation of complex argument, it is generally agreed that all too often it is utilized inappropriately by both teachers and external examiners.

This trend of scant attention being paid to 'less important', everyday classroom assessment, with greater efforts being made for 'important' and 'competitive' external examinations in terms of validity and reliability, seems to be universal and unlikely to be reversed in the foreseeable future.

Matters of usability (time to administer and mark a test, ease of administration, cost factors, presentation, etc.) do not appear to be of great concern to teachers in most countries, probably through oversight, rather than deliberate neglect.

Teachers are apparently wedded to traditional types of assessment for particular purposes, regardless of shortcomings. For example, it is clear that pencil-and-paper tests, often of an objective type, are resorted to in order to assess instructional outcomes that would be more validly measured by, say, practical or oral tests. But the impracticality and cost-effectiveness of such techniques mitigates against their use and the loss of validity must therefore be tolerated. Until more efficient and effective techniques (of proven validity and reliability) are devised, many of the desired indicators of learning and achievement will be inadequately and inappropriately assessed in most, if not all participating countries.

Aspects of test presentation and formatting, which affect face validity and student motivation, also require much closer teacher attention.

Standardized tests of achievement, where they are employed by teachers in the broad assessment of learning, or are used for purposes of moderation, e.g., in Australia, Japan, New Zealand and the Philippines, are usually of known high reliability and high to moderate validity, and are designed to be attractive in terms of usability, many of them being machine-scorable.

In summary, it is clear that the basic concepts underpinning relevant (valid) and trustworthy, consistent (reliable) educational assessment, as one component of student evaluation, are not fully recognized or are poorly understood. And the essential techniques required to ensure sound measurement are practised infrequently by the vast majority of teachers and examiners in most, if not all countries. Consequently, irrelevant and unreliable information is too often incorporated in decision making for individuals and groups in terms of students' learning progress and achievement.

Regrettably, the situation is unlikely to change markedly in the short term. Change in a positive direction is dependent on the provision of appropriate courses in measurement and evaluation in both the pre-service and in-service training of teachers; something that is sadly lacking in most of the countries participating in the seminar.

2.6 Grading and Marking

Marks are commonly given by both teachers and examiners in all countries, and these are usually converted to percentages for ease of understanding.

Some examining bodies apply scaling to raw marks. Aspects of scaling for externally examined subjects, particularly hierarchical scaling where it is carried out, as in Australia and New Zealand, are highly controversial when the statistical treatments become publicly known, even though scaling is done to try to be fair to all students. Many countries use an A-E letter grading system in the assessment of student achievement, although numerical scales are also used.

2.7 Effects of Assessment

Most of the participating countries provided information on the effects of assessment on the larger community and on the students, teachers and the entire instructional system.

There is an increasing concern over the use of tests, particularly aptitude measures, in perpetuating the power and privilege of the dominant culture, or the middle-class within a country's society, and in discrimination against ethnic or cultural minorities, usually associated with allegations of test bias (New Zealand and the Philippines).

Failure in the National College Entrance Examination (NCEE) in the Philippines, for example, is a stigma that high school graduates seek to escape. Hence, they keep taking the NCEE in order to pass. Those who fail after several attempts are relegated to vocational and technical courses that take less than 4 years. Similarly, in Australia, Japan and New Zealand, opponents of examinations argue that they have harmful effects, particularly tests which permit 'invidious comparisons between students' and the built-in failure of a proportion of them. The tendency in these situations is to move away from external examinations towards more subjective forms of assessment that are school-based.

In India, on the other hand, the trend has been towards the establishment of a National Testing Service to provide quality-controlled tests and to rationalize the giving of entrance tests required by several different institutions, specifically those offering medicine, engineering, business administration, etc.

The negative effects of assessment are commonly considered to outweigh the positive ones. However, Malaysia and Sri Lanka explicitly expressed the motivational aspect as a positive effect of assessment on students. The perceived, but not necessarily demonstrated, negative effects are summarized as follows:

1. On Students:

(a) The preparation for and taking of tests and examinations causes undue anxiety, stress or tension in many students (Australia, China, India, Indonesia, Japan, Malaysia, New Zealand and Sri Lanka).

(b) Students resort to devious means of passing examinations, e.g., mass copying, bribing the examiners, leakage of examination papers, etc. (India).

(c) The types of questions used in assessment, most of which are short-answer or essay type, force students to resort to 'memorization rather than trying to understand' (India, Nepal and Pakistan).

2. On Teachers and Teaching Procedures:

While possible negative effects of assessment on the student were not observed in Indonesia, Nepal, the Philippines and Thailand, special mention was made that teachers tended to be more concerned with covering the curriculum rather than taking up all curricular offerings. Their goal was to prepare students for the end-of-year external examinations, but thereby sacrificing quality. As Sri Lanka further reports, 'teaching becomes examination-oriented', and thus a gap is created between the intended curriculum and the achieved curriculum, a situation also encountered in Japan and the Republic of Korea. In China and Indonesia, it is observed that most teachers tend to teach their students to concentrate only on the kinds of things on which they will be tested, hence teachers emphasize the teaching of factual information. Formative evaluation is believed to be grossly neglected, to ensure adequate time to cover examination syllabuses (India and Sri Lanka).

3. On the Instructional System:

Mention has also been made of the effect of assessment on the entire instructional programme. It has been observed in Japan that there is an overload of the whole learning programme due to very frequent testing. Test results are also used for ability grouping of students, a practice which parents and students have severely criticized. In Sri Lanka, teaching programmes in the upper primary grades are reported as 'distorted' due to the external Grade 5 scholarship examination which acts to defeat the objectives of the official primary curriculum. Another effect found in Sri Lanka is where the progress of rural schools has been retarded because pupils who excel in the scholarship examination have left to go to urban schools.

Society at large and school authorities have been discussing the pros and cons of assessment. This, of course, is nothing new. The main objection centres on the over-use of objective, paper-and-pencil tests. The desired change is to increase the use of performance measures and tests in the affective domain and to ensure other non-cognitive measurement. In essence, teachers and examiners are required to broaden the range of their assessment and to utilize all the techniques available to them.

All countries reported on internal and external examinations. One of the more controversial issues was the content of external examinations which may not have been covered in instruction or assessed in the specific teacher-made tests. Most of the anxieties of students and teachers arise from this lack of matching, especially if the external examination is considered as the basis for promotion and streaming students in the different categories of courses in college/university. One possible way to improve the situation is to have a proper mix of internal and external examinations, as in the trend specifically mentioned by Indonesia and Malaysia and senior secondary school examinations in Australia and New Zealand.

2.8 Teachers' Competencies in Assessment

Ten out of thirteen countries reported that teachers generally have inadequate training in the development of assessment tools and in making systematic, relevant observations. Specific competencies which teachers need to develop are:

(a) Preparation of diagnostic and summative assessments;

(b) Development of criterion-referenced and norm-referenced tests;

(c) Thorough knowledge of evaluation methodologies;

(d) Construction of items congruent with stated instructional objectives;

(e) Construction of application and of understanding test items;

(f) Carrying out formative evaluation as the monitoring component of a goal-oriented instructional model for mastery learning;

(g) How to select and administer appropriate assessments;

(h) Analysis and interpretation of test results;

(i) Utilization of test results.

All the countries that cited teachers' inadequacies and a lack of thorough training in assessment, also reported that plans for improved in-service training were available, but that in most instances these had not been fully implemented. In Pakistan a large scale programme of teacher training in evaluation has been carried out since 1974, although the level of training is not very high. The general observations were that the present situation is unlikely to improve. If there is any change, it will probably be very slow because of (a) prevailing negative attitudes towards assessment, (b) a lack of teacher motivation and/or recognition of the problems, and (c) a non-conducive environment in many teaching-training institutions.

Recommendation

The locus of the problems addressed in this section of the report, for all countries, is clearly the individual teacher and his/her competencies in the procedures and techniques of valid and reliable educational assessment.

It is recommended that:

Adequate human and financial resources be devoted to the improvement and enhancement of teachers' skills and understanding in the sensitive and sensible uses of educational assessment at both pre-service and in-service levels.

3. DOMAINS OF ASSESSMENT

3.1 Cognitive Domain

The educational objectives of almost all countries include the total development of the child, including the cognitive, affective and psychomotor domains. However in practice, the education system in each country tends to over-emphasize the cognitive domain and under-emphasize the psychomotor and affective domains, particularly the affective domain. This becomes very obvious where public examinations are used for assessment.

Most countries have reported that assessments tend to emphasize cognitive objectives (Australia, India, Indonesia, Japan, Malaysia, Nepal, New Zealand, Pakistan, the Republic of Korea and Sri Lanka). Thailand reported that all tests cover cognitive and psychomotor domains well, only the affective domain is under-represented. The Philippines reported this problem is common, especially with teacher-made tests.

Many countries also reported that among the objectives in the cognitive domain, there is a preponderance of items which assess lower-level cognitive objectives, and very few items assessing at the higher-level (India, Indonesia, Nepal, Pakistan and Sri Lanka). Malaysia and the Republic of Korea have reported this problem with teacher-made tests only.

There are many reasons for this over-emphasis on the cognitive domain. While the policy statements of the educational authorities emphasize the overall growth and development of the individual as the major goal of education, the cognitive domain is seen as the major preserve of the schools. Further, more direct and objective measures for assessment of cognitive objectives are available. Teachers also have greater expertise in constructing tests for cognitive objectives.

3.2 Psychomotor Domain

Curricula in all countries involve some aspects of psychomotor development. Many countries include training in physical education, games and sports (Australia, China, India, Indonesia, New Zealand, Nepal, Pakistan, the Philippines and the Republic of Korea) as a part of school work and report the performance of the student to the parents. However these reports are based on informal and sporadic observations, except probably in those cases where excellence has been demonstrated in competitive sports. There is usually no external examination for such activities and excellence in them does not count in assessments as determined by public examinations.

Assessment in the psychomotor domain is carried out internally for certain subjects such as Music, Art, Engineering and Technical Drawing (Australia and New Zealand), practical skills (Indonesia), Science subjects (Australia and Pakistan), and Home Economics (Australia, Pakistan and Sri Lanka). Practical tests are used for all appropriate subjects in Malaysia. Assessment in the psychomotor domain is also done internally in China. In the Philippines some teachers conduct individual and small group tests in certain Physical Education skills, Practical Arts, Home Economics, Industrial Arts, Music and Art. In Australia, it is common that teachers of subjects involving a practical component provide a statement that a satisfactory practical course has been completed by each student, as a requirement of the Examining Board.

The major problems in assessment of the psychomotor domain, apart from those mentioned earlier, are the large numbers of candidates in external examinations, which makes practical examinations where manipulative skills can be assessed/observed very difficult to administer.

3.3 Affective Domain

The affective domain seems to experience the most problems. Some countries, such as Australia and New Zealand perceive formal moral assessment as incongruent with their secular education systems. Educationists in the Republic of Korea feel that affective assessment may constitute a violation of the privacy of individuals. Assessment of the affective domain does not constitute a part of formal public examination in any country.

There is a great deal of concern among all the participating countries regarding development of affective aspects of personality. The concerns include development of pro-social behaviour (politeness, punctuality etc.,) and positive attitudes towards school and learning. Some informal and subjective assessment is made by teachers in some countries such as Malaysia. In the Philippines teachers sometimes use situation tests in subjects including Health Education, Social Studies, Work Education and Home Economics. Tests for the measurement of Study Attitude, and academic self-concept have been developed in some countries (Pakistan), but these are not being widely used. New Zealand primary and intermediate school teachers are required to rate students on a norm-referenced scale from A-E on traits of stability, independence, perseverance and cooperation.

Apart from the problems of ethics and ideology, the major constraint in the use of affective assessment is methodological. It is extremely difficult to obtain information regarding the affective domain with a degree of reliability and validity that would justify its inclusion in public or even internal examinations.

Most countries express a need for monitoring affective development of the child in order to develop pro-social behaviour and proper attitudes towards learning. Many countries expect that the direction of change to restore a semblance of balance would be towards greater importance/emphasis for the affective and psychomotor domains (Australia, India, Indonesia, Japan, Malaysia, Nepal, New Zealand, Pakistan and the Republic of Korea). As a solution, New Zealand proposes to officially change its assessment practices to provide more qualitative comment with a concomitant emphasis on less-structured and possibly less reliable, but more comprehensive, data gathering. This is also being done in Australia, parallel with the development of reliable instruments to measure students' attitudes.

Recommendations

It is recommended that:

1. There is a need for incorporating items of higher level objectives of the cognitive domain in internal evaluation as well as public examinations.

2. Psychomotor and affective domains have been by and large neglected in all forms of assessment in educational institutions, with quite undesirable feedback on educational practices. There is a need for more research in this area, and development of better assessment instruments. However, even at the present level of development they should form a part of internal assessment by the teachers.

4. EQUITY AND EXCELLENCE

4.1 Quality and Quantity

All countries of the region accept the responsibility of providing some education to all their citizens. Providing equal opportunity of education to all citizens and attaining and maintaining a high level of excellence are also accepted goals of all these countries.

Many countries of the region have a system of pre-primary education. They include Australia, Japan, New Zealand and Sri Lanka in which 70% to 80% of children are enrolled. Fewer students have access to pre-primary education in China, India, Indonesia, Pakistan, Nepal, the Philippines, the Republic of Korea and Thailand.

A number of countries such as Australia, China, Indonesia, Japan, Malaysia, New Zealand, the Philippines, the Republic of Korea, Sri Lanka and Thailand have either universal enrolment in primary education or are very close to it (above 90 percent). Other countries such as India, Nepal and Pakistan are still moving in this direction.

A particular problem in India, Indonesia, Pakistan, Nepal, the Philippines and Thailand is a high rate of drop-out. These countries follow a system of annual promotion examinations. In other countries (Australia, Japan, Malaysia, New Zealand, the Republic of Korea and Sri Lanka), students move automatically from one grade to another with their age peers.

All the countries are concerned about the quality of education. Regular surveys for monitoring the system are carried out only in China, Indonesia, Malaysia and Sri Lanka, although surveys on an irregular basis have been carried out in all the countries of the region. These latter surveys are usually at regional level but many countries have also carried out national surveys. Some of them (Australia, India, Indonesia, Japan, New Zealand, the Philippines, the Republic of Korea and Thailand) have also participated in the studies of educational evaluation conducted by IEA (International Association for the Evaluation of Educational Achievement). In Nepal, studies of the achievement levels of children completing primary school have been carried out on a regional and national level in 1980 and 1984 by CERID (Centre for Educational Research, Innovation and Development).

4.2 Success and Retention/Retrieval

Most countries of the region have a system of linkage so that those who have dropped out and want to return to secondary or higher education can either return to the main stream or obtain further education through distance learning techniques.

A person may elect to return to school to complete his/her secondary education in Australia and New Zealand. In both these countries any adult, who while at school did not meet the requirements of university admission, may enter under provisional matriculation. In Indonesia and Pakistan an Open University caters for the needs of students who cannot enrol in a college or university. In India such students are provided with opportunities for continuing their studies through correspondence courses in an open school. The use of radio and television through satellite technology is also increasing. In the Philippines over-aged drop-outs or school leavers and those who have not been to school at all may take the Philippine Educational Placement Tests and return to the formal educational system. Education from primary to university level is available to adults in China. The open university teaching is carried out using radio and TV. Trade schools and adult centres are available in Nepal for over-aged students or drop-outs from school.

4.3 Disadvantaged Groups

Many countries have identified disadvantaged groups in their population. Such groups are females (Australia, China, India, Nepal, New Zealand and Pakistan); rural groups (China, India, Indonesia, Malaysia, Nepal, Pakistan, the Philippines and Sri Lanka); ethnic minorities (Australia, China, India, Nepal, New Zealand and the Philippines), and people inhabiting remote areas (Australia, China, India, Indonesia, Nepal, New Zealand, Pakistan, Philippines, Sri Lanka and Thailand).

A number of schemes have been undertaken to overcome these problems. In Australia, schools with a significant proportion of aborigines and students from other non-English-speaking backgrounds receive additional funds for staff and facilities. Similar schemes are available in New Zealand. In Indonesia and the Philippines large regional differences exist in the level of students performance. Use of local norms for national examinations in such cases is acceptable, especially in student certification.

In India and Pakistan the bulk of non-enroled and drop-out children belong to disadvantaged sections of their societies. Special schemes and innovative projects have been sponsored by the federal governments of these countries or by international agencies to overcome this problem. Students in remote and border areas of China are also disadvantaged. In Nepal remote regions pose special problems for which major innovative projects have been initiated. Although 85% of secondary students stay on to Form 5, Maoris and Pacific Islanders and low socio-economic students in New Zealand tend to leave school as soon as they are legally able to do so, but this tendency is slowly being reversed. The Philippines offers special scholarships to minority communities. Students from such communities are encouraged to return and serve their areas.

Recommendations

It is recommended that:

1. There is a need for development of diagnostic tests for remedial education.

2. National and international studies of educational achievement serve a useful purpose in raising the quality of education. Regional collaboration whenever possible should be encouraged.

3. There is a need for more research on the performance of disadvantaged groups on tests of scholastic aptitude and achievement. Additional funding for scholarships, staff and facilities, along with some measures of positive discrimination for such groups are recommended.

5 MONITORING LEVELS OF ACHIEVEMENT

5.1 Purpose of Monitoring

The overall purpose of monitoring levels of achievement is to improve both the quality and quantity of student education. Within that overall purpose, the more immediate concern is to obtain the best possible information about student performance. Once this information is obtained, it becomes possible to identify areas of need and to plan programmes for the improvement of student learning. There is always a need to improve the quality of the information obtained about students, as student assessment information can only be as good as the instruments used in obtaining the data.

Although different types of studies are discussed in this section of the report, the monitoring studies described here all have the major intention of improving the system of education available to students. Measures of achievement made for other purposes, such as the certification of students, have been considered earlier in this report.

5.2 Types of Monitoring

There are two types of monitoring which were reported by the participating countries, and which are included in this section of the report. These are the monitoring of the examinations themselves, and the monitoring of student achievement. In some countries both types of monitoring are carried out, in others only one type, and some countries do not engage in either of these types of monitoring studies.

Examination Monitoring: All participating countries recognize the importance of improving the assessments undertaken, especially when these are made on a regular basis. In particular, Indonesia, Japan, Malaysia, New Zealand and Sri Lanka emphasize the monitoring of their annual examinations with the intentions of maintaining and improving reliability and validity. In these cases, the annual examination is used to monitor the educational system as well as for other purposes. Although all the other countries also do some work on assessment monitoring, this activity assumes lesser importance because the instruments concerned are not used as part of the regular examination system. India and Pakistan, in particular, stated that they do not have regular monitoring studies of this type.

Monitoring Student Achievement: In some participating countries, student achievement is assessed by means of monitoring studies conducted by an agency external to the school, usually the Education Department or other body responsible for schooling. One other example of monitoring studies is the IEA surveys of achievement in reading, mathematics, science and written composition. In most countries student achievement is monitored regularly, but such assessments are more limited or irregular in Australia, Nepal and the Philippines. In other countries, notably New Zealand, the schools do their own monitoring and results are not made available publicly on a more general basis at regional or national levels, although the data are available to the central government department.

5.3 Mastery and Minimum Learning

At least at the primary (or elementary) level of schooling, all participating countries expect their students to achieve a minimum level of performance in basic skills, especially in language and arithmetic. For example, in China the minimum level of language performance for students in rural areas is that they should be able to read and write at least 2,500 Chinese characters by the end of primary education. Requirements of students in urban areas of China are higher.

In some participating countries these minimum skills are listed in some detail in curriculum statements and, in some cases such as New Zealand, are referred to as the core curriculum. Other countries, particularly Australia, have not specified basic skill curriculum requirements, but students are expected to reach levels of performance in the basic skills which will enable them to succeed in their later schooling, and to have the capability of living a normal, full and productive life in the community after leaving school. Most frequently, it is the basic skills or core curriculum areas that are monitored by educational systems, in those countries where external monitoring takes place. About half the participating countries (Australia, India, Japan, Malaysia, Nepal, New Zealand and the Republic of Korea) reported that no external monitoring was done on a regular basis. There was no general trend with respect to external monitoring - some countries favoured increasing monitoring, while others were reducing it.

There has been some adoption of the concepts of mastery learning into curricula and the assessment of students. This is particularly true in the Republic of Korea in which students are expected to achieve a mastery level of approximately 80 percent in all subjects studied at primary and secondary schools. Also the Philippines has set a target of 75 percent performance for all students in external examinations. In Indonesia, the concepts of mastery learning are being developed in the Development School Projects through the use of learning packages at the primary and secondary levels. School curricula in Australia do not reflect mastery learning concepts, but the two national assessments of basic skills that have been undertaken have required an 80 percent mastery level for success on specific tasks. It is possible that interest in mastery learning will increase in other participating countries, and that other applications will develop.

Assessments are also made of higher-level learning in secondary schools for the purposes of investigating and improving aspects of the system of education. The IEA surveys, referred to above, are examples of studies in which the full range of learning within particular subject areas is assessed. Most of the participating countries have been involved in at least one IEA survey of student achievement.

A major issue which most countries face with respect to student achievement monitoring concerns the possible backwash effects on students of the assessments made. In particular, concern is expressed that the concentration on basic skills in most monitoring studies could cause distortion of the curriculum such that too great an emphasis is given to minimum competence. The opposing opinion is that the assessments provide an incentive for schools, teachers and students to achieve the required basic levels. The potential positive and negative effects on students of competition in external examinations is also relevant here. Research is needed into the effects on students of all external assessments.

5.4 Uses of Research on Assessment

The use made of research on assessment, including surveys of student achievement, should be sufficient to justify the time and expense of the research. Consequently, simple surveys of student achievement, which do not also gather data on student and family background and school information (including the curriculum being taught), are probably of limited value. It is important for researchers to be able to understand and explain the strengths and weaknesses of student performance and the differences in achievement between various groups of students. All participating countries agreed that, where possible, major monitoring studies on student achievement should include other explanatory information to assist in interpreting the results obtained. In particular, information on the curriculum actually taught should be obtained and reported.

It is also important that the staff of the schools participating in research on achievement accept that the information being collected is likely to be of benefit to them in either the short or the longer term. It is the professional responsibility of those conducting such research to ensure, to the maximum extent possible, that the studies done are useful to the participants. As schools in a number of countries are becoming increasingly autonómous, it is probable that it will become necessary, as well as desirable, to convince the principal and staff that the intended research will be beneficial, or the schools may decline to cooperate in the study.

A significant difference in the nature of the research on assessment between the participating countries is the level at which it is conducted. In some cases the assessment is national, in others regional, and in others local (which includes the work of individual university researchers). Most assessment for the purpose of monitoring achievement is done at the national level in Indonesia, Philippines and the Republic of Korea. In addition, at least some assessments are national in those countries which participate in the IEA surveys. Regional level surveys predominate in Australia, China, Japan, Philippines and Thailand. Research and surveys of students' achievement are conducted at regional or national level in Nepal, whereas this work tends to be more locally based in India, New Zealand and Pakistan.

A major use of research on assessment in most countries is for curriculum development. The assessments may reveal areas in which students have low levels of achievement, and which may require greater attention by teachers. Those responsible for the curriculum should use the achievement information obtained to help in deciding whether they should increase curriculum emphasis in the areas where achievement was low.

Another significant use of assessment research is the provision of detailed achievement information to schools on their students' results. These data can be used in giving remedial and other assistance to individual students where necessary. Of course, such feedback to schools must be provided quickly if it is to be of use, and local research or research in very small systems is best placed to facilitate such fast feedback.

Other uses, each mentioned by at least one country, include the improvement of teaching methods, the development of objectives, and educational planning of facilities. Details of the types of research on assessment and the uses to which the research is put in the participating countries are contained in the individual country reports found in Part II of this report.

Recommendations

It is recommended that:

1. A greater use of bilateral contacts and group seminars on specific aspects of assessment should be made. Examples of specific aspects of assessment worthy of study include: (a) how to deal with the competitive aspects of external assessments; and (b) the levels of student performance considered to represent minimal competency in the various countries of the Asian and Pacific Region (which differ considerably), and the methods by which the levels are determined.

2. An emphasis should be given to research on those aspects of assessment likely to be most useful to educational systems and to schools. For example, the research should include additional information on the student and the school to enable a greater understanding and utilization of assessment results.

3. Given the increasing independence of schools in some countries, greater attention must be given to making research on assessment more immediately useful to the schools involved in the study.

SUMMARY OF RECOMMENDATIONS

1. If assessment is accepted as a multitrait-multidimensional process of gathering data, then its scope needs to be enlarged in respect of all countries, particularly where assessment is too narrowly conceived. (Section 1)

2. In the countries where external examinations dominate the curriculum, there is a need to make more use of formative assessment, while simultaneously decentralizing the process of assessment to the teacher, and the use of a variety of informal techniques. (Section 1)

3. There is a need to have more comprehensive information, particularly concerning student aptitudes and interests, for guidance purposes. (Section 1)

4. It is considered desirable that certificates issued should contain more information than is generally provided at present on certificates. (Section 1)

5. Many countries recommend that national monitoring of educational achievement should be implemented, but this recommendation was not unanimous. (Section 1)

6. There is a need to give proper stimulation to accelerating the level of attainment of the brilliant and accomplished students in the cognitive and non-cognitive areas. (Section 1)

7. More balanced assessment, including information in the affective domain, should be made, where possible, to reduce an over-emphasis on cognitive achievements. (Section 1)

8. Adequate human and financial resources should be devoted to the improvement and enhancement of teachers' skills and understanding in the sensitive and sensible uses of educational assessment at both pre-service and in-service levels. (Section 2)

9. There is a need to incorporate items of higher level objectives in the cognitive domain in internal assessment as well as in public examinations. (Section 3)

10. There is a need for more research into all forms of assessment, particularly in the psychomotor and affective domains. Even at the present level of development, psychomotor and affective assessment should form part of internal assessments made by teachers. (Section 3)

11. There is a need for development of diagnostic tests for remedial education. (Section 4)

12. National and international studies of educational achievement serve a useful purpose in raising the quality of education. Regional collaboration, wherever possible, should be encouraged. (Section 4)

13. There is a need for more research on the performance of disadvantaged groups on tests of scholastic aptitude and achievement. Additional funding for scholarships, staff and facilities, along with some measures of positive discrimination for such groups are recommended. (Section 4)

14. A greater use of bilateral contacts and group seminars on specific aspects of assessment should be made. (Section 5)

15. An emphasis should be given to research on those aspects of assessment likely to be most useful to educational systems and to schools. (Section 5)

16. Given the increasing independence of schools in some countries, greater attention must be given to making research on assessment more immediately useful to the schools involved in the study. (Section 5)

CONCLUDING REMARKS

In the same way that education cannot be considered in isolation from the wider society, the assessment of educational achievement should not be divorced from the context of education as a whole. Although the linkages of assessment with the broader educational context have been mentioned only briefly in Part I of this report, the individual country reports in Part II provide the necessary, more detailed information on the general educational contexts in which the various assessments are embedded.

Many of the issues and problems in the assessment domain were found to be shared by all the participating countries. Both the common problems and the different solutions are reflected in this report. The sharing of information about how the various countries viewed the problems and how they were attempting to solve them was, of course, most valuable. As is so often the case, it was found that we can all learn from one another in the area of assessment, as in other matters.

The coverage of this seminar on assessment was necessarily very wide, and the papers and discussions covered many different types of school examinations and research on assessment. The seminar provided a good foundation for further sharing of ideas in this area. However, because the topic was so broad, it would be desirable to hold subsequent seminars on some of the more important aspects of assessment, which could then be treated in far greater depth. Examples of some specific aspects that could be studied are contained in the recommendations.

Part II

INDIVIDUAL COUNTRY REPORTS

The Assessment of Educational Achievement in Australia

S. F. Bourke

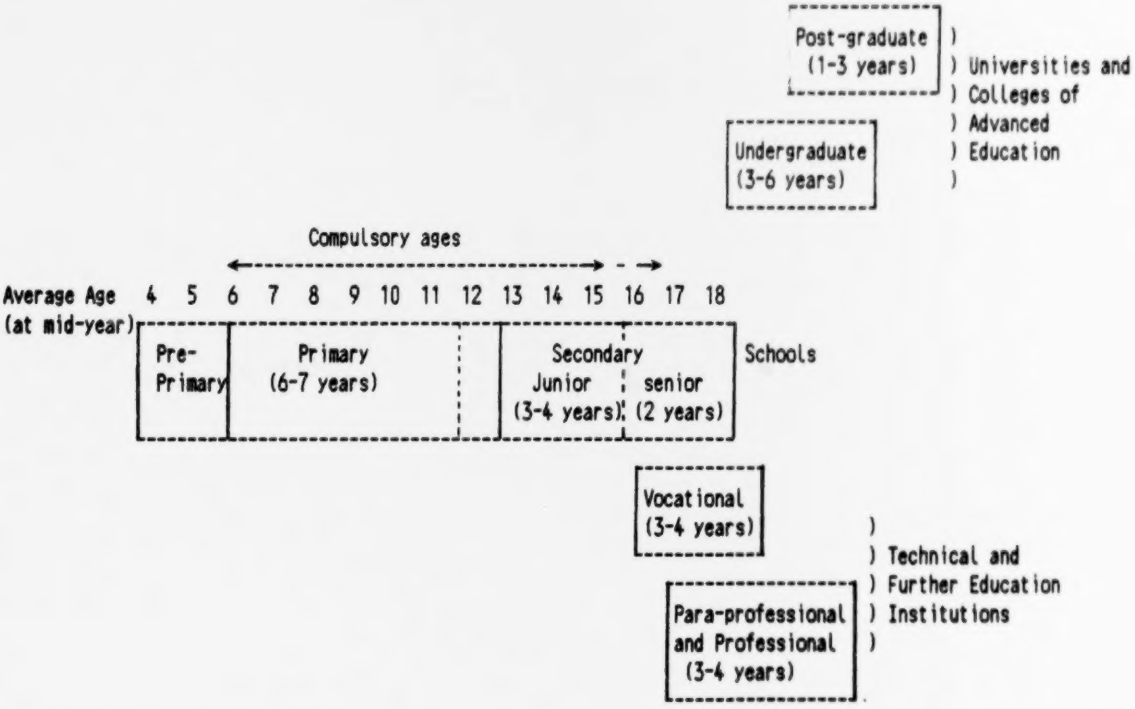
AUSTRALIAN EDUCATION SYSTEMS

Under the Constitution, education in Australia is a State rather than a Federal responsibility. Consequently, no national system of education exists. Although the various State systems do exhibit structural similarities, there are curricular differences between systems and a variation in patterns of assessment which apply in the six States and two Territories of Australia. The Federal and State Ministers of Education meet regularly as the Australian Education Council but, although some decisions are made, the Council is largely a national forum for the exchange of views on educational priorities and policy issues.

The overall structure of education in Australia is shown in Figure 1. In general, there are 12 years formal schooling, made up of either six years of primary and six years of secondary education, or seven years of primary and five years of secondary education. Pre-primary education, where provided, is not compulsory. The compulsory ages of attendance are from six to 15 years, except for Tasmania where attendance is compulsory until 16 years of age. Secondary schools are comprehensive (except for some technical schools in Victoria) and usually coeducational. Except for Tasmania and the Australian Capital Territory (ACT), secondary schools span the full five or six years. In Tasmania and the ACT, students in the final two years (Years 11 and 12) are segregated into secondary colleges. As a result of the high per student costs involved in Years 11 and 12, the concept of secondary colleges is being discussed in other states, notably in Victoria where the establishment of such colleges has been recommended to the government. Secondary colleges have been set up on an experimental basis in Victoria and Queensland.

Across Australia, 66 per cent of all secondary students completed Year 11 in 1984, while others moved into vocational education and the workforce. About 45 per cent completed Year 12, which represents a large increase in retentivity since 1980. Of those enrolled in Year 12 in 1983, 41 per cent went on to higher education in 1984 (Ainley, Batten and Miller, 1985:6), the remainder going into vocational education (usually part time) and the workforce.

About a quarter of all Australian students attend non-government schools, most of which have a religious affiliation.


Post-graduate
(1-3 years)
Universities and
Colleges of
Advanced
Education
Undergraduate
(3-6 years)
Compulsory ages
Average Age
(at mid-year)
4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Pre-
Primary
Primary
(6-7 years)
Secondary
Junior
(3-4 years)
senior
(2 years)
Schools
Vocational
(3-4 years)
Technical and
Further Education
Institutions
Para-professional
and Professional
(3-4 years)

(Source: P. Mckenzie "Australia: System of Education", International Encyclopedia of Education, 1985: 370.)

Figure 1. The Structure of Australian Education

These schools require registration by the State, and receive Federal and State Government financial support on both a per student and a needs basis.

In addition to differences between State systems, there are increasing differences within systems as the State Departments of Education, to greater or lesser extents, devolve administrative and curriculum responsibilities to regional offices and to individual schools. As far as curriculum decision-making is concerned, devolution of responsibility to schools has been particularly marked in some states at the junior and middle secondary levels of education. More generally, schools and individual teachers now have a large measure of autonomy in the structuring of teaching and for the selection of learning materials (Bourke, 1984:105).

In spite of the State responsibility for education, the role of the Federal (or Commonwealth) Government in education has increased markedly over the last 20 years as the States have become increasingly financially dependent on the Federal Government. Approximately 50 per cent of the general purpose grants from the Federal Government to the States is spent on education (McKenzie, 1985:372). In addition to the general purpose grant, the Federal Government also provides specific purpose educational grants to the States. The increasing interest of the Federal Government in education is relevant here because it has led to decisions to implement surveys of educational achievement which would not have been taken in other circumstances.

STUDENT ASSESSMENT

Two types of assessment of educational achievement, with quite different purposes, are distinguished in this paper.

1. *Student Assessment* for the purpose of deciding each individual student's progress to the next stage of education or providing each student with a recognized qualification.

2. *System Monitoring* for the purpose of providing information on aggregate student performance, for appropriate groupings of students, to those responsible for the educational systems, or to the schools, or both. These assessments may be made of age or grade samples of students, and information fed back to schools may be used to assist individual students. However, assessments made for monitoring purposes are not used in making decisions about student progress or certification for individual students.

Student assessment is considered now, and system monitoring is discussed in the next section.

PRIMARY AND JUNIOR SECONDARY ASSESSMENT

Within the primary and junior secondary school systems in Australia, progression is largely a function of age, with students moving through the grade levels with their age peers. This is particularly true of primary schools. Severely educationally retarded students may be retained in a grade, but this is a rate practice even in those states where it does happen. Repeating a year is more common in the junior secondary years, but again is not often suggested. The increasing tendency is to place physically and mentally handicapped students into normal classes with their age peers, in so far as this is practicable.

No external tests or examinations are used during primary and junior secondary education. Especially in primary schools, teachers generally make informal assessments of students, sometimes coupled with standardized tests which they choose to use. In junior secondary schools, assessments are frequently based on projects, assignments, class cooperation, teacher-designed tests, some use of standardized tests, and internally-conducted examinations in some states. There is some ability streaming of students in the junior secondary years in some states, but in other states the practice is officially discouraged by the departments of education and, as a result, is virtually non-existent.

About one third of each age group leaves full time secondary education at the end of junior secondary schooling (Year 10), although this proportion is declining. These students normally obtain a statement of school-based assessment of subjects studied which may take the form of a certificate provided by the State Department of Education.

SENIOR SECONDARY ASSESSMENT

The assessment of students at the conclusion of their secondary education is required to serve two purposes: first a summary of achievement across a range of subjects studied which may be used as a certification that the student has achieved a level of performance in his or her studies; and secondly a comparison of where each student stands in relation to his or her peers who sat for the same examination, which is used in the selection of students where demand for places exceeds supply, either in the workforce or in further education. There are a number of potential conflicts in the dual purposes served by end-of-school assessments, among them general questions of breadth versus depth in the curriculum, and the related issue of the extent of any core of subjects which should be studied by all. In resolving such conflicts it is important that at least the dual purposes of Year 12 assessments are recognized. In a study of certification in the upper secondary school, McGaw and Hannan found no reason why Year 12 certificates should not serve divergent purposes, but they considered that at least the certificates should be designed to minimize conflicts between the purposes (McGaw and Hannan, 1985:7-9). In supporting the value of the two purposes described above - affirmation of end-of-school achievement and selection of candidates - McGaw and Hannan warned that over-emphasis on the latter could undermine the former, and argued that striking a balance between the purposes is perhaps the most urgent issue facing upper secondary school assessment in Australia.

Table 1: Certification Procedures for Years 11 and 12 in Australia by State and Territory

State/ Territory	External Exams (Yes/No)	Exam/School Contribution	Scaling Used (Yes/No)	Minimum No. of Subjects	Aggregate Provided (Yes/No)	School-based Courses Moderated (Yes/No)	Certificate available in Year 11 (Yes/No)
New South Wales	Yes	50/50	Yes (Proposed:No)	Usually 4 (based on units)	Yes (Proposed:No)	No (Proposed:Yes)	No (Proposed: Yes)
Victoria	Yes (Proposed:Some subjects only)	50-70/50-30 (Proposed: 50/50)	Yes (Proposed:No)	4	Yes	No (Proposed:Yes)	No (Proposed: Yes)

Table 1: Certification Procedures for Years 11 and 12 in Australia by State and Territory (continued)

State/ Territory	External Exams (Yes/No)	Exam/School Contribution	Scaling Used (Yes/No)	Minimum No. of Subjects	Aggregate Provided (Yes/No)	School-based Courses Moderated (Yes/No)	Certificate available in Year 11 (Yes/No)
Queensland	No (since 1973) (Criterion-referenced from 1986)	School only	Yes (using ASAT)	3 (major subjects)	Yes	Yes	No
South Australia	Yes (May be reduced or dropped	75/25 (Proposed: 50/50	Yes (Proposed:No)	5	Yes (Proposed:No)	Yes	Yes
Western Australia	Yes (plus attainment of literacy from 1986)	50/50	Yes (using ASAT)	5 (at Year 12 level)	Yes	Yes	Yes
Tasmania	Yes	50-75/50-25 (since 1981)	Yes (within subjects only)	4 (major)+ 2 (minor)	No	Yes	Yes (It is possible to enter university after Year 11)
Australian Capital Territory	No	School only	Yes (within subjects only)	3 (major)+ 1 (minor)	Yes	Yes (Using ASAT)	No (but since 1985 a "Record" has been available)
Northern Territory	Yes (South Australian exam is used)	75/25	Yes	5 (as for South Australia)	Yes	Yes	Yes

Each State and Territory of Australia has its own formal system of assessing each individual student's educational achievement at the conclusion of secondary school (Year 12). With the exceptions of the State of Queensland and the Australian Capital Territory, the majority of students' assessments are based in part on externally set and marked examinations. Certification in Queensland and the ACT is based on school assessments with some adjustment of scores intended to remove potential inequities between schools in tertiary selection. Schools in the other five states and in the Northern Territory also offer school-based courses which are assessed internally, usually with some moderation of results between schools. Although such courses cater for a minority of students completing secondary school, the proportion is increasing in most states. A recent study by Ainley, Batten and Miller (1985:14) suggested that the increasing proportion of students in school-based courses rather than in the more traditional external courses was in part due to increases in student retentivity by schools where school-based courses were offered. An effect on the proportion of students staying on to Year 11 was particularly noticeable. The suggestion was based on earlier study in which the same researchers found that, after allowing for social background influences, retentivity to Year 12 in the State of Victoria was significantly greater in schools which offered courses which were alternatives to the traditional externally examined courses (Ainley, Batten and Miller, 1985:15). It would seem that the provision of school-based courses is helping to retain many students who had previously left school before the completion of their secondary education.

A brief summary of certification procedures in Years 11 and 12 across Australia is provided below. States are listed in descending order of population, as by McGaw and Hannan, from whom these summaries have been extracted (McGaw and Hannan, 1985: Appendix B). Major points have been highlighted in Table 1. As there are many changes pending or proposed for certification in the states, the more significant changes are also indicated in the table.

New South Wales (NSW)

The external examination at the end of Year 12 in NSW is known as the Higher School Certificate. A proposal exists for the introduction of a Certificate of Secondary Education for all students who do not complete Year 12. As with most systems using external examinations, the results of the examinations are combined with school-based assessments for the same courses to provide an overall score for each subject (or each unit of study). Subjects may be studied at a number of levels so, since 1983, a common test has been used to scale the results for the different groups of students within a subject. In NSW, scaling is currently used between subjects as well as within subjects, although a proposal exists to abandon the former.

The number of subjects used in selecting students for higher education varies depending on the levels of subjects studied. If the proposal to no longer provide an aggregate of scaled scores for tertiary entrance is adopted, tertiary institutions may decide to impose subject prerequisites to ensure that potential studies are compared on the basis of similar patterns of subjects (McGaw and Hannan, 1985:66).

Victoria

The external Year 12 examination in Victoria is known as the

Higher School Certificate, although there is a proposal to replace all existing certificates with a single one to be known as the Victorian Certificate of Education. A distinguishing feature of the upper secondary years in Victoria is the proliferation of programmes available. In addition to the courses leading to an external examination and the school-based subjects accredited by the Victorian Institute of Secondary Education (VISE), there are two other major programmes which have varying degrees of acceptability for different purposes. Established in 1976, the Schools Year Twelve and Tertiary Entrance Certificate (STC) has, since 1981, also been accredited by VISE and has the aim of facilitating the participation of teachers and students in the development and assessment of courses. Within technical schools, similar Year 12 courses have been developed, and are collectively known as Tertiary Orientation Programmes (TOP). In each case the course is negotiated between teachers and students, and frequently developed in conjunction with particular tertiary institutions to ensure acceptability for entry to specific courses.

Students seeking university entrance normally take five subjects, although four subjects is the minimum requirement. Most tertiary institutions recognize the STC and TOP courses in some way for admission, although aggregates derived from these courses may be scaled down against aggregates derived from the Higher School Certificate courses. A proposed broadening of upper secondary school courses is currently being discussed with a view to reducing the importance of the aggregate score in selection. It is also proposed that the external examination be held in the middle of Year 12 rather than at the end of that year.

Queensland

A Senior Certificate is awarded to students completing Year 12 wholly on the basis of a school assessment. Moderation of results between schools is achieved through District and State review panels which consider student work and other supporting documents submitted by schools, and offer advice to the schools on their proposed achievement level awards. A number of reviews of this scheme, implemented in 1973, pointed to several difficulties. Student-teacher relationships had suffered because the teacher had become the students' judge as well, in a competitive situation. The amount of testing increased as teachers sought evidence on which to base their assessment decisions, and the testing frequently focussed on very specific content areas. As a result of these and other problems, the system of school-based assessment was reviewed and a new system will be fully implemented in 1986. Assessments will be based on defined performance criteria which have been written into courses in the form of objectives. Student achievement will be measured against specified objectives rather than against the achievement of other students. As McGaw and Hannan stated:

> This new system represents the first serious attempt in

> Australia at a system-wide level to reject normative assessment and embrace criterion-referenced assessment. (McGaw and Hannan, 1985:70)

At the end of Year 12, students receive a separate Tertiary Entrance Statement providing an aggregate of their performance over Years 11 and 12. Subject assessments are scaled for all students in each subject within each school using the Australian Scholastic Aptitude Test (ASAT), which all students applying for tertiary entrance must take during Year 12. This test provides a general aptitude measure and more subject specific, though not syllabus based, measures similar to the U.S. College Board's test. (Test specifications for ASAT were published by the Australian Council for Educational Research in 1978.) An aggregate is obtained by adding adjusted scores for subjects, and then adjusted for the school as a whole, again using ASAT.

South Australia

Two certificates have been available to Year 12 students in South Australia depending upon the course followed: a Public Examinations Board Certificate and a Secondary School Certificate, the latter offered by the Education Department. These have now been combined into a single certificate. Examination results and school assessments are scaled and then combined to give subject results. School-based assessments are moderated through school visits and meetings when work and assessments are reviewed. The Chief Moderator for a subject may change teachers' assessments. A range of options concerning external examinations is being canvassed for the period after 1986. It seems likely that the weighting given to the external examination component of subject results will be reduced, if the external examination is retained.

University entrance is based exclusively on subjects examined externally, although the colleges of advanced education take some account of school assessed and moderated subjects, after discounting results against the externally examined subjects.

Western Australia

From 1986, the Certificate of Secondary Education is the only certificate available to Year 11 and 12 students. The external examination results are combined with a school assessment moderated against the external examination results to give the subject results. Results are scaled using the Australian Scholastic Aptitude Test. Student achievement for Years 9 to 12 are shown on the certificate, although scaled marks are provided only for Year 12 results.

Students applying for tertiary entrance from 1987 will be required to meet three criteria:

1. a Secondary School Graduation requirement which includes

achievement of a satisfactory literacy level,

2. a satisfactory performance in six subjects in Year 12, with at least five subjects at the Year 12 level,

3 an admissions average calculated from the best three, four or five subjects, depending on subject area.

Tasmania

The Higher School Certificate is available to students completing Years 11 or 12, showing results in both externally assessed and school assessed subjects. School assessments are moderated against the external examination results where possible, and otherwise through visits to schools by advisers who approve syllabus coverage and methods of assessment. The whole question of upper secondary courses and assessment is currently being discussed, including the possibility of using criterion-referenced assessment similar to the Queensland model.

Since 1982, university entrants have required passes or lower passes in a total of six subjects. No aggregate score is provided. University admission is still possible after Year 11, although it is less frequent than previously. College entrance requires only four subjects, but completion of Year 12 is necessary.

Australian Capital Territory (ACT)

From 1986, students who complete at least five of the six terms of Years 11 and 12 will be awarded the ACT Year 12 Certificate. An ACT Secondary College Record is available to students who leave school before then. Assessment of the Certificate is internal to the school, and subjects are scaled.

Supplementary Information is provided for students seeking tertiary entrance. Subject scores are adjusted using the Australian Scholastic Aptitude Test as a moderator. An aggregate of each student's three best scaled scores plus 0.6 of the next best major or minor score is calculated.

Northern Territory

Students receive a Northern Territory Senior Secondary Studies Certificate after completing at least one-semester course in Years 11 and 12. The certificate is based on courses and assessments made by the Senior Secondary Assessment Board of South Australia and school assessments which are moderated on the South Australian results. Consequently, those students completing Year 12 also receive a South Australian certificate.

As the Northern Territory does not have a full range of tertiary institutions, the majority of tertiary aspirants seek entrance to institutions in other states on the basis of their

South Australian assessment.

OVERVIEW OF STUDENT ASSESSMENT TRENDS IN AUSTRALIA

Since the 1960's, the various Australian education system have abandoned all external assessments previously made during primary and junior secondary schooling. More recently, all systems have also moved away from external examinations in the upper secondary school (Years 11 and 12), including assessments made at the end of Year 12. Although an end-of-school external examination still exists in six of the eight systems, in no case is it the sole criterion of assessment. If current proposals are accepted, in most states the external examination will account for no more than half the total assessment with school-based assessment providing the other half. In two of the systems there is no external examination, although both use the Australian Scholastic Aptitude Test (ASAT) to scale subjects or to moderate results for tertiary entrance. If other State systems were to cease conducting external examinations without using ASAT in conjunction with school assessments, considerable care and effort would need to be invested in moderating results between schools, in order that selection for tertiary education or employment would be equitable.

There is also the possibility that moves away from the traditional normative assessment methods, in which students are ranked to provide results, will gain momentum at the Year 12 level. Criterion-referenced assessment, based on student performance in meeting specified objectives is used in Queensland for Year 12 results, and similar possibilities have been recommended in recent years by examining bodies in Victoria and Tasmania. If these or other State systems feel able to overcome the difinitional and measurement problems associated with criterion-referenced assessment, the major criticisms of normative assessment as "ensuring a proportion of failures" and "making invidious comparisons between students" will no longer be relevant.

A final major issue of curriculum and assessment in Years 11 and 12 is the balance to be found between the specific requirements of tertiary entrance qualifications and the more general desirability of a rounded education. Whereas the former has been dominant in the past when a high proportion of students in he upper secondary schools had the intention of undertaking some form of tertiary education, the latter now assumes increasing importance as higher proportions of students stay on to complete Year 12 without intending further study. Two alternative solutions would seem to be:

1. Accept a bifurcation of courses and assessments depending on each student's intentions. Major problems here would be the certainty of status and acceptability difficulties for non-tertiary entrance courses, and resource problems for schools attempting to run two separate courses.

2. Insist on a general education for all to Year 12, and thus delay specialization for two years. Major problems here would be convincing the tertiary institutions that at least one preparatory year should be added to their present undergraduate courses, and the impatience of many upper secondary students, who have decided what they want to do, to begin specialization in their chosen subject areas.

No doubt other options will also be canvassed as this relatively new and important problem is faced. With all eight education system in Australia in a state of flux, solutions adopted to this and other curriculum and assessment problems in the upper secondary school may well differ.

SYSTEM MONITORING

In the last decade a number of surveys have been conducted of student achievement in primary and secondary schools. The surveys, some national and others regional, have assessed different subject matter at various age or grade levels. Only the major surveys are summarized here.

NATIONAL SURVEYS

Two types of national surveys have been conducted in Australia: studies of literacy, numeracy and oracy (listening and speaking), and studies conducted in cooperation with the International Association for the Evaluation of Educational Achievement (IEA). The IEA studies carried out in Australia have concentrated on mathematics and the physical sciences.

All of these national surveys have assessed the achievement of students by age or by grade at two or three selected levels. within each State and Territory, schools were sampled with a probability proportional to the number of students in the target population, and the same number of students (usually 25) were tested at each school selected. Consequently each student in the State or Territory had an equal chance of being selected in the sample. State and Territory samples were then weighted according to their target populations and national estimates of achievement were calculated, usually in the form of the proportions of students correct on test items and mean scores on groups of items. In each case, background information on the students assessed was also collected, and results were reported for various groups of students.

Studies of Literacy and Numeracy

Surveys of reading, writing and numeration were conducted in 1975 and 1980. The first study, commissioned by a Federal Parliamentary Committee, indicated an increasing Federal interest in education. Following this survey, there were several inquiries into education which recommended that some form of assessment programme in the basic skills should be implemented (see Bourke, Mills, Stanyon and Holzer, 1981:1-3). This was the climate in which the second survey was commissioned by the Australian Education Council (AEC) which is made up of the Ministers of Education, State and Federal. Both the 1975 and the 1980 surveys were carried out by the Australian Council for Educational Research. A summary of the 1980 study follows.

Background to the Study. The Australian Studies in Student Performance (ASSP) project consisted of a survey of literacy and numeracy skills of two samples of more than 5000 students aged 10 and 14 years in normal schooling throughout Australia. Assessments were made of the proportions of students who could perform a range of tasks, and it was possible for all students to be considered successful on a given task. The assessment of students' abilities in reading, writing and numeration was designed to enable statements to be made concerning the proportions of students at these two age levels who were 'at risk' in terms of normal functioning at school and in the wider community because of inadequate skills in these basic skill areas. The information obtained could also be compared with similar information collected in 1975. (The 1975 data were presented by Bourke and Keeves, 1977.)

Results. The proportions of students who were successful in performing the various tasks were reported separately by content area, by skill and by type of usage (classroom or social). An overall proportion of successful students for the reading and numeration tests each taken as a whole was also provided. (See the report by Bourke, Mills, Stanyon and Holzer, 1981.) An overview of some of the results is presented in Table 2, in which the percentages of successful and unsuccessful students at ages 10 and 14 years are provided for the total reading and numeration tests, and for two of the writing tasks: writing a personal message within the family, and writing a more formal letter. As an extension of the results obtained, estimates of the approximate numbers of student throughout Australia who could be expected to be unsuccessful on these tests and who would therefore be in need of help in at least some of the areas tested, are also given in this table.* It can be observed from the table that in every case there was a higher proportion of successful students at age 14 years when compared with age 10 years. This was the case even though the tasks tested were normally more difficult, but were considered suitable for the students' level of development. When the potential numbers of students 'at risk', and thus requiring further help are considered, it would seem that between 57,000 and 75,000 students aged 10 years throughout Australia would be in this category for reading and numeration, while the number of similar students for writing depended very much on the nature of the task involved. In the case of the 14 year-old students, it would seem that up to about 40,000 students could be having problems with reading and numeration, and fewer students were troubled by writing, but again this very much depended on the task.

---

* Evidence available from the 1975 ASSP work suggested that there was a strong relationship between success on the reading and numeration tests as the correlation between total test scores was in the range of 0.6 to 0.7 (Bourke and Keeves, 1977:104). Consequently the majority of the students needing help in one area would also need help in the other.

Table 2: Overview of ASSP Results: 1980‡

Area	% Successful		% Not Successful		Number Needing Help	
	Age 10	14	10	14	10	14
Reading	71	83	29	17	75,000	41,000
Numeration	78	86	22	14	57,000	34,000
Writing:						
A personal message	93	97	7	3	18,000	7,000
A letter	57	87	43	13	112,000	32,000

‡ The percentage in this table have been extracted from several tables in Bourke et. al. (1981:70, 74, 92, 98, 118 and 128).

Comparison with 1975. When the results for the tests administered in 1975 and 1980 were compared, at both age levels, the students either had the same or a higher level of performance in 1980, depending on the nature of the task. In particular:

1. when reading a prose passage approximately five per cent more students aged 14 years were successful in 1980;
2. when reading a newspaper, 26 per cent more of the 10 year-olds, and 11 per cent more of the 14 year-olds were successful in 1980;
3. when performing computation using the four arithmetic operations, there was an improvement in 1980 on three-quarters of the items by the 10 year-olds, and on almost half the items by the 14 year-olds. The average improvement was approximately five per cent for the 10 year-olds and four per cent for the 14 year-olds;
4. when reading maps and scales, calculating volume and using percentages, there was an average improvement in 1980 for the 14 year-olds by a margin of six per cent over 1975;
5. when writing a personal message, seven per cent more of the 10 year-olds were successful in 1980; and

6. when writing a more formal letter, 27 per cent more of the 10 year-olds wrote a greeting at the beginning of the letter and nine per cent more wrote their name at the end.

On all other tasks tested, there was no significant difference in performance between 1975 and 1980. In no case at either age level did performance decline between 1975 and 1980.

It was contended that the development of literacy and numeracy was a pre-requisite for many aspects of further learning and for a normal life in the contemporary Australian community. In almost every assessment made, the majority of students were successful, and for many tasks the proportion of successful students was high. While some reactions to the results of this study have focused on this satisfactory result, others have been concerned with the educational and social consequences for the smaller numbers of students who were unable to do some of the tasks.

Postscript to the 1980 Study. The Australian Education Council (AEC) had originally intended that the 1980 survey was to be the first of a series of similar surveys in order to map student performance in literacy and numeracy over time. However, this plan was opposed by some teacher unions, some academics and, in a more muted way, by some officers of State Education Departments. What was seen as a political, Federal initiative was not universally welcomed. For a summary of the arguments concerning assessments such as these see Marsh (1982). The 1980 survey was evaluated by a team of academics which concluded that, although the study had been properly conducted, 'few, if any' state policy or curriculum changes seemed likely to result from the study (Power, 1982:114). As a result of the opposition referred to and of the evaluation report, in 1983 the AEC finally abandoned intentions for further national surveys of literacy and numeracy.

The Oracy Study (Listening and Speaking)

A study of listening and speaking in Australian schools was carried out from 1976 to 1979, between the two literacy and numeracy surveys described in the previous section. The oral language curriculum was, and still is, an area largely neglected in Australian education, and little information was available on school practices and assessments. Consequently, the study was carried out in three phases which are now described briefly.

Literature and Curriculum Survey. This first phase of the study provided an overview of the concept of oracy and its relationship with literacy, in and out of school. Clark and Davis described the lack of information on student performance in listening and speaking as 'remarkable' given the importance of oral language, and suggested that the lack of information on students' oral language development was a 'major breakdown' in Australian education (Bourke, Clark, Davis and Holzer, 1980:26-27).

Investigating Teaching Practices and Teacher Opinion. Lists of listening and speaking tasks which could be included in an oral language development programme were prepared, and a framework for classifying the tasks was suggested by Bourke and Holzer. The framework consisted of three dimensions: the ability students needed to do the task, the context in which the task was performed, and the purpose served for the students in listening or speaking. Group interviews were then conducted with teachers at primary and secondary schools to ask them which tasks they considered to be important, what they actually did in assisting students to develop oral language abilities, and how they assessed listening and speaking. Although there were some differences between primary and secondary teachers, they generally agreed on the important tasks. Finally, lists of 22 listening and 22 speaking tasks selected for possible assessment were provided (Bourke et al., 1980:90).

Assessment of Listening and Speaking. Based on the lists of important listening and speaking tasks identified in the second phase of this study, tests were developed to assess the listening and speaking performances of 10 year-old and 14 year-old students throughout Australia. The listening tests were recorded on tape to provide standardized administration, and the speaking tests were conducted as interviews and recorded on tape for subsequent scoring. Approximately 1300 students at more than 200 schools were tested at each age level using mainly criterion-referenced assessments. Results were reported by the proportions of successful students on each listening and speaking task tested, supplemented by an overall listening assessment. A selection of the results is shown in Table 3 for both the listening and speaking tests. It will be noted that there was a wide variation in performance for the different tasks, indicating the importance of discrete assessments of this type in identifying problem areas. As expected, the performance of the 14 year-olds was generally higher than the 10 year-olds when the same task was assessed at the two age levels.

Table 3: Some Listening and Speaking Test Results: (a)

Task(b)	Successful Students (%) Age 10 years	Successful Students (%) Age 14 years
Listening Tasks		
* Comprehending words and statements	97	100
Understanding statements	100	100
* Comprehending passages	25	74
Comprehending conversation	65	88
* Comprehending in different situations	47	80
Total Listening Test	62	79

Table 3: Some Listening and Speaking Test Results:(a)-continued-

Task(b)	Successful Students (%) Age 10 years	 Age 14 years
Speaking Tasks		
* Giving personal details	80	95
Giving directions	72	64
* Repeating a story	87	92
* Repeating an idea	46	72
* Making critical judgments	63	84
* Giving own ideas	85	95
* Giving a message	80	71
Reading aloud	96	94

(a) Extracted from Bourke et al. (1980:127, 147).

(b) An asterisk indicates a task that was identical at both age levels.

Conclusions. The identification of specific tasks and competencies for both teaching and assessment purposes was considered to be the most practical and useful way to proceed for oracy as had been done for literacy and numeracy (see Keeves and Bourke, 1976). Of course not all teachers accept the notion of language development based on specific abilities, although many do use these ideas (either explicitly or implicitly) in oral language development programme. At the very least, most would agree that oracy deserves a more prominent place in school curricula. In a subsequent summary of the study, it was suggested that oracy (listening and speaking) should be considered with literacy and numeracy as basic skills which form the basis for all further learning and which all students should possess on leaving school (Bourke, 1980:3).

IEA Studies of Mathematics

The IEA studies of mathematics in 1964 and 1978 were carried out in Australia by the Australian Council for Educational Research. The latter study is summarized here. The mathematics achievement of students at two levels was assessed: namely students aged 13 years and students in Year 12 who were studying mathematics. The tests were based on an analysis of the mathematics curricula over several countries prior to testing in 1964. The same tests were used in 1978, when it was claimed that the tests were probably more closely linked to mathematics curriculum content in Australia than they had been in 1964 (Rosier, 1981:6). Student attitudes to mathematics and more general attitudes were also assessed. The summaries of results that follow have been extracted from the full report of the study (Rosier, 1980).

Results for 13 year-old Students. There was a slight overall decline across Australia in mathematics achievement between 1964 and 1978 as measured by a 65-item test. In 1978, mean scores ranged from 23.1 in Victoria to 29.0 in Queensland. However, Victoria was the only State where the decline in mean scores was statistically significant at the 0.05 level. When various areas of mathematics were considered separately, there was a lower level of geometry achievement in 1978, but a higher level of achievement on items related to 'new' mathematics. Such a result might be expected given changes in emphasis in mathematics curricula with decreasing time given to geometry and increasing time to 'new' mathematics across Australia between 1964 and 1978.

Results for Year 12 Students. The State mean scores on a mathematics test of 69 items were higher in 1978 than they had been in 1964, although the difference was significant only in Western Australia. In 1978, mean scores ranged from 23.3 in Western Australia to 33.1 in Tasmania. The higher scores in 1978 were remarkable given that there had been an increase of approximately 69 per cent in the proportion of students remaining at school until Year 12 and in the proportion of students studying mathematics at this level.

Results for Student Attitudes. There had been an overall deterioration in the attitudes of students towards mathematics and more generally, at both age levels between 1964 and 1978. Specific attitudes, all of which showed a decline, were:

1. the importance of mathematics for future employment,
2. the importance of mathematics for understanding the environment,
3. perceptions of the ability of people to control their physical and social environment, and
4. that schools and school work were enjoyable.

Conclusion. This study stressed the importance of curriculum aspects of studies of achievement, in particular when comparisons in achievement over time were made between different groups of students. It was suggested that varying emphasis on mathematics, in terms of time spent in class on teaching mathematics in the various Australian states, may have contributed to the differences in achievement observed. Rosier also argued that the observed decreases in mathematics performance of the 13 year-old students between 1964 and 1978 were partly due to the replacement of carefully formulated course specifications in 1964 with more general guidelines for teachers in 1978. However, Rosier went on to suggest that the general improvement in mathematics performance at the Year 12 level indicated that the curriculum at that level should not be a cause for concern, but that the changing nature of the population of Year 12 students might require significant changes to mathematics curriculum offerings in the near future (Rosier, 1980:198-200).

Perhaps the most disquieting aspect of the results of this study relate to the decline in students' attitudes towards mathematics and school. This occurred despite reductions in class size, increased levels of teacher qualifications and, at least for the Year 12 students, higher levels of mathematics performance. The establishment of a balance between emphasis on cognitive achievement and student attitudes is worthy of some consideration.

IEA Science Studies

National surveys of science achievement were conducted by the Australian Council for Educational Research in 1970 and 1983-4. The 1970 survey focused on the achievement and attitudes of 14 year-old students and students in Year 12 (see Rosier, 1973 for curriculum and assessment details). The 1983-4 survey was concerned with science curriculum, achievement and attitudes of primary students aged 10 years, secondary students aged 14 years, and Year 12 students.

A preliminary report of the 1983-4 study was prepared for the APEID Regional Meeting on 'Science for All' at Bangkok in 1983 (Rosier and Kay, 1983). Australian results for the study should become available in 1986, with various international reports being prepared from 1986 to 1988.

REGIONAL SURVEYS

A number of surveys have been carried out periodically by State Education Departments during the last 10 to 15 years, notably in New South Wales, Queensland and Tasmania. These studies are summarized below. Some very recent work on identifying objectives and developing assessments has also been carried out in the Northern Territory. No report of these results was available in time to include a summary in this paper.

New South Wales (NSW)

The Central Metropolitan Region of the NSW Department of Education conducted a survey of literacy and numeracy skills of students in government schools in Years 5 and 8 in 1976. Detailed feedback to schools was provided on the reading, written expression and numeracy tests used. School follow-up and region-wide activities were instituted in an attempt to assist schools and students in need. It was not suggested that this Region was representative of all students in NSW.

The overall results for the Region were reported by Nay

(1976). A norm-referenced approach to assessment was used in that the results were grouped into five categories, and students in the two lowest categories were considered to have 'a level of skill inadequate for adult life'. Whereas only small percentages of students at both year levels were considered to be unsuccessful in written expression, much higher proportions of students were identified as needing help in numeracy and reading. This was particularly true for reading in which about four-fifths of the Year 5 and almost half the Year 8 students were considered to have an inadequate level of skill.

Queensland

The Queensland Department of Education conducted reading surveys of Year 5 students on three occasions between 1971 and 1981, and study skills surveys of Year 7 students in 1972 and 1978. These surveys are now briefly summarized.

The Reading Surveys 1971-81. The purpose of this work was to compare reading achievement of Year 5 students in Queensland government primary schools in 1981, with reading achievement of Year 5 students in 1971 and 1976. The results were reported by McBryde, Phillips and Byrne (1982) in terms of mean scores and differences between means on the three occasions. No information was provided on the absolute levels of performance on the test in terms of the proportions of successful and unsuccessful students, or on the nature of the tasks tested. There was no statistically significant difference in reading comprehension (implied, literal or total), and there were decreases in speed and vocabulary since 1971, although not between 1976 and 1981.

The Study Skills Surveys 1972, 1978. The Queensland Department of Education administered four tests of study skills to samples of Year 7 students in government primary schools in 1972 and 1978. The results of Map Reading, Graph Reading, Index and Dictionary Tests have been reported by Duck (1980), again in terms of mean scores and differences between means. Criteria of successful completion of the tests were not established. There was no change in achievement between 1972 and 1978 in the Graph Reading, Index and Dictionary Tests, but a slight decline on the Map Reading Test. However the validity of four items to which the decline was largely attributable was questioned in the report (Duck, 1980:vii).

Tasmania

The Tasmanian Education Department used the reading and numeration tests developed by the Australian Council for Educational Research in 1975 for the first national study of literacy and numeracy to test all their students aged 10 and 14 years in 1975. Subsequently they have progressively modified the tests and assessed these age groups at various time intervals up to the present. The proportions of the total groups of students in three categories, have been reported periodically. (See for example, Broadby, 1980). In the 1980 reports, the three categories were referred to as 'competent' students (answering at least 90 per cent of the items correctly), students for whom the results were 'inconclusive' (between 75 and 90 per cent), and students 'at risk' (less than 75 per cent).

In 1979, the proportions of students 'at risk' ranged between about 30 and 40 per cent. However, the basis on which the categorization of students was made was quite arbitrary, and was not related to the intentions of the original test developers who constructed tests in which approximately 80 per cent of correct items indicated success, or to the likelihood of error when classifying students into the three categories. Where comparable data were available over time, Broadby (1980b:5) suggested that there had been a significant improvement in reading skills of the 14 year-old students between 1977 and 1979. Over that period an additional 10 per cent of students were classified as competent, and 4 per cent fewer students were considered to be at risk.

One important aspect of the Tasmanian work in this area is the identification of schools which had higher proportions of students 'at risk', with the intention of providing greater assistance to those schools. In this way any additional help available could be targeted to the areas of greatest need.

OVERVIEW OF SYSTEM MONITORING IN AUSTRALIA

The national surveys of literacy, numeracy and oracy assessed student achievement in terms of the proportions of students who could perform specific tasks considered to be important for further progress through the school system, or in the students' everyday lives. Thus the assessments made were criterion-referenced, and students who could not perform the tasks were considered to be 'at risk'. This type of assessment was considered to be most suitable for measures of student performance in basic skills, although the selection of tasks to be assessed required careful consideration.

The various regional surveys summarized made use of different assessment instruments, and frequently identified the students with the lowest performance as being 'at risk', simply because they were in the lowest categories. These assessments were normative, with the certainty that some students had to fall into the lowest categories. As different assessment gradings were used in these surveys, it is not possible to compare or combine their results in any way.

If the detailed results of the 1980 national survey of literacy and numeracy are taken as indicative of the proportion of students 'at risk', it would seem that between 20 and 30 per cent of the 10 year-old students were in this category, and between 10 and 20 per cent of the 14 year-old students, both proportions depending on the specific tasks assessed. Translating these proportions of 'at risk' students into classes of 30 students, an average of between 5 and 8 students in the upper primary school, and between 3 and 5 junior secondary students in each class were in need of at least some additional help with basic skills development.

OTHER NATIONAL ASSESSMENT INITIATIVES

A major criticism of surveys of educational achievement is that they do not of themselves directly assist student learning. Such surveys provide basic information of a national accounting type which, when allied with other student, school and community information enables specific groups, schools and areas to be targeted for additional help. However, these arguments are frequently not accepted as sufficient justification for the expense of the surveys. Other objections relate to the focus on basic skills which was a feature of the surveys carried out. Concern is expressed that the surveys may distort the curriculum towards too great an emphasis on basic skills at the expense of other learning. Two other national initiatives related to assessment are worthy of specific mention.

AUSTRALIAN COOPERATIVE ASSESSMENT PROJECT

With a widespread shift to school level responsibility for curriculum and assessment and for school evaluation, there have been accompanying moves to assist schools to undertake these tasks. In particular the Australian Cooperative Assessment Project, set up in 1979 and given increased emphasis in 1983 when national assessment was formally abandoned by the Australian Education Council, has the aim of providing assessment materials for schools which would enable them to undertake their own, more professional assessments of student learning and of the curriculum offered. Unfortunately little cooperative work has proved to be practicable to date, but at least personnel in the various State and Territory Education Departments now have this mechanism for making the materials they develop available to the other Departments.

QUALITY OF EDUCATION REVIEW COMMITTEE

In 1984, a different type of initiative was the establishment by the Federal Minister for Education of a Quality of Education Review Committee. The Committee, consisting of academics and government officers, was charged with the responsibility of examining the effectiveness of Federal involvement in primary and secondary education to assist the Government in developing clear, more effective strategies for directing its funds for school level education. One priority area investigated by the Committee was the identification, assessment and development of basic skills, with a view to improving the quality of education in these areas. The Committee reported to the Minister in April 1985 (Karmel, 1985).

A significant task of the Committee was to consider the means by which the effectiveness of Federal involvement in education should be assessed. Although the Committee did suggest that more stress should be placed on student outcomes, and that evidence of student attainment of general competences should be provided every three years, it did not propose that a regular programme of national assessment should be undertaken for this purpose (Karmel, 1985:107). The Committee emphasized the importance of teacher education, both pre-service and in-service, in achieving the necessary improvements.

FOR THE FUTURE

An overall impression of education in Australia is the rapid changes taking place. Two associated areas seem to be at the forefront of change: the increasing responsibility of schools, and changes to curriculum and assessment in the upper secondary school. Whereas basic skills and teaching methods may be relatively constant, the locus of responsibility for educational decision-making is clearly moving to schools in most States. In some cases schools are demanding a greater role in curriculum and planning to meet their perceptions of the needs of their students, in other cases schools are having these responsibilities thrust upon them as a result of Education Department policy decisions.

The assessment of educational achievement in Australia seems certain to continue the current movement towards school-level responsibility, both for student assessment and certification and for monitoring the school curriculum. As a result of this movement, it is necessary that teachers receive sufficient help through consultancy and in-service education to enable them to undertake these tasks competently and confidently. Many schools are currently required by their State Education Department to undertake their own school evaluations. These schools frequently seek outside advice on what types of cognitive and affective outcomes should be assessed, and what means should be used. In general, it would seem that non-test information will be increasingly used for student and school assessment.

Changes to the types of assessment made are also likely to continue. Informal, teacher-made tests are implicitly criterion-referenced in nature, although usually not explicitly, because they are designed to test what the teacher has taught, whether this is content or a skill. The tests also provide diagnostic information for individual students, although this is frequently neglected by teachers. Teachers need help in fully utilizing assessment information, particularly for diagnostic purposes. The further development of moves to criterion-referenced assessment in the more formal examination systems would remove one of the major criticisms of traditional testing: that it is competitive and consequently requires that a proportion of students fail. Whether criterion-referenced assessment will prove to be acceptable in Australia as a basis for selecting students for tertiary education, which is increasingly competitive, remains to be seen. Perhaps the establishment of more senior secondary colleges would concentrate teacher expertise in these smaller institutions and make it possible to obtain more complete information on tertiary aspirants for use in selection.

Student assessment in primary and junior secondary school has established in recent years, with individual schools having a clear responsibility. The greater role of school-based assessment at the Year 12 level is part of the increasing role of schools in decision-making. When this is combined with a flow of major structural changes to the external examination system, where it exists, and an apparent shortage of university places for successful Year 12 students, it can be seen that the situation is complex. Once the current and proposed changes to Year 12 assessment have been implemented, there should be sufficient variation between the Australian States and Territories to enable them to observe how alternative systems operate, without needing to experiment themselves. All may benefit from a few years of relative stability in assessment procedures in the upper secondary school.

REFERENCES

Ainley, J., Batten, M. and Miller, H. Patterns of Retention in
1985 Australian Government Schools: Summary Report. Canberra: Australian Government Publishing Service.

Australian Council for Educational Research. Research Papers
1978 Relating to the Australian Scholastic Aptitude Test. Hawthorn, Victoria: Australian Council for Educational Research.

Bourke, S.F. "The Assessment of Oracy in Australian Schools."
1980 ACER Newsletter No. 40.

Bourke, S.F. The Teaching and Learning of Mathematics. (ACER
1984 Research Monograph No. 25). Hawthorn, Victoria: Australian Council for Educational Research.

Bourke, S.F., Clark, M.L., Davis, D.F. and Holzer, F. Oracy in
1980 Australian Schools. (ACER Research Monograph No. 9). Hawthorn, Victoria: Australian Council for Educational Research.

Bourke, S.F. and Keeves, J.P. (Eds). The Mastery of Literacy and
1977 Numeracy. Canberra: Australian Government Publishing Service.

Bourke, S.F., Mills, J.M., Stanyon, J. and Holzer, F. Performance
1981 in Literacy and Numeracy: 1980. Canberra: Australian Government Publishing Service.

Broadby, F.M. 1979 Survey of Basic Number Skills of 10-Year-Old
1980a Tasmanian Students. Research Study No. 51. Tasmania: Education Department.

Broadby, F.M. 1979 Survey of Basic Reading Skills of 14-Year-Old
1980b Tasmanian Students. Research Study No. 52. Tasmania: Education Department.

Duck, G. Achievement Levels of Queensland Year 7 Pupils in
1980 Certain Study Skills. Queensland: Department of Education Research Branch.

Karmel, P.H. (Chairman). Quality of Education in Australia.
1985 (Report of the Review Committee). Canberra: Australian Government Publishing Service.

Keeves, J.P. and Bourke, S.F. Literacy and Numeracy in Australian
1976 Schools. Volume 1: A First Report. Canberra: Australian Government Publishing Service.

McBryde, B., Phillips, K. and Bryne, M. A Comparison of the Read-
1982 ing Achievement of Queensland Year 5 Pupils Between 1971, 1976 and 1981. Queensland: Department of Education Research Branch.

McGaw, B. and Hannan, B. Certification in Upper Secondary Educa-
1985 tion. Canberra: Curriculum Development Centre.

McKenzie, P. "Australia: System of Education." The International
1985 Encyclopedia of Education. (Editors T. Husen and T.N. Postlethwaite) Oxford, England: Pergamon Press. pp. 368-76.

Marsh, C. "Point and Counterpoint. National Testing and the
1982 Curriculum: Suspect or Saviour?" Curriculum Perspectives 2(1): 77-88.

Nay, W. (Chairman). Enquiry into Literacy and Numeracy. Report of
1976 the Committee for Enquiry into Literacy and Numeracy. New South Wales: Department of Education.

Power, C. (Coordinator). National Assessment in Australia: An
1982 Evaluation of the Australian Studies in Student Performance Project. Canberra: Australian Government Publishing Service.

Rosier, M.J. Changes in Secondary School Mathematics in Australia
1980 1964-78. (ACER Research Monograph No. 8). Hawthorn: Australian Council for Educational Research.

Rosier, M.J. "Changes in Secondary School Mathematics in Aust-
1981 ralia 1964-78." ACER Newsletter No. 41.

Rosier, M.J., and Kay, C.M. A Report on the Position of Science
1983 Education in Australia. Hawthorn, Victoria: Australian
Council for Educational Research. (mimeo).

On Standards for Investigating the Achievement of Universalization of Primary Education in China

Wang Yibing

The universalization of primary education (UPE) is a prerequisite for the development of a civilized society and also a foundation for training qualified personnel for modernization. As early as the end of the last century, the Qing Government of China at one time raised the issue of UPE. Subsequently the Guomingdang Government many times designed plans for UPE. But until 1949, the enrolment rate of school age children had only reached 25%. Of the whole population, 80% were illiterate and illiteracy was as high as 95% in the countryside. Since 1949, the Chinese Communist Party and the Chinese Government have given much attention to the development of education. More than 30 years later, the enrolment rate of school age children reached 95% in 1984 and the percentage of illiterates among the people in 1982 was reduced to 23.5%. In 1980, the Chinese Government decided that UPE will be realized except in remote and isolated areas. In order to facilitate the historic task of UPE in a developing country like China with a population of one billion, the Ministry of Education in China (it was changed into State Commission of Education in June, 1985) in 1983 issued the basic requirements for investigating the achievement of UPE.

I. The Basic Requirements for UPE

In view of the diversity of physical conditions in China and the uneven development in the economy, culture and education, the Ministry of Education put forward the basic requirements for UPE at the present time:

1. The enrolment rate of school age children should be 95%.

This refers to the percentage the children of school age in full-time schooling and attending simple primary schools of various types including teaching classes or groups occupy in the whole of children of school age in China. The children of school age in the areas where the five-grade teaching system is practised refer to children between seven and eleven, and to children between seven and twelve in the areas where six-grade teaching system is practised. At present retarded children are mostly at primary schools and are included in the calculation except those who are seriously retarded and do not have the ability to learn. Deaf and dumb children are separately calculated.

2. The annual retention rate of children at school is maintained at over 97%.

This refers to the comparison of the number of children at the beginning of a term and that at the end of an academic year. The calculation method is like this: the number of children at the beginning of the second semester plus the number of graduating children, minus the number of the intakes at the beginning of a new academic year.

3. The rate of graduating children in the graduating grade. In the cities the figure should be 95%, it should be 90% in the countryside where the condition is fairly good, and for the rest of the areas 80%.

This refers to the actual number of graduates in the total of children in the graduating grade. The graduating children are given an all-round examination: morally, intellectually and physically. Apart from these graduates, the children in the graduating grade who failed in the graduating examination but who can recognize 2,500 commonly used Chinese characters are considered to have acquired the basic ability of reading and writing, and can perform the four fundamental operations of whole number of arithmetic, are also granted certificates of completion of the programme. Those failures continue to study at school. The graduation examination at the countryside primary schools are under the guidance of county educational departments and the test papers are designed by the township leading school according to the basic requirements of the teaching plan and are prepared in several versions for different schools. The graduation examination in the cities is decided by the educational departments of various provinces, cities and autonomous regions.

4. The rate of UPE among the children between twelve and fifteen should reach 95%.

This refers to the rate of graduates among children between twelve and fifteen in the total of this age group. In the countryside, especially in the areas where the economic and educational conditions are poor, those children who, in various ways, can recognize 2,500 commonly used Chinese characters and can perform the fundamental operations of whole number of arithmetic are also included in the calculation of the rate of UPE. At the same time, the number of those children who continue to study at school are deducted from the total of this age group and those who have gone on to secondary schools are included in the group.

II. The General Method of Investigating the Achievement of UPE

The responsibility for achieving UPE was under the leadership of the governments of the provinces, autonomous regions and municipalities directly under the central government, and their respective educational departments are in charge of the specific work.

The general method is as follows:

1. According to the constitution, the requirements of the central government, and on the basis of the actual situation in different provinces, autonomous regions and the municipalities, the legislative bodies of the provinces, autonomous regions and the municipalities formulate the laws and regulations of the UPE in the locality. There are generally detailed regulations concerning the leadership and management of UPE, for instance, the obligation of the governments at all levels, the obligations of parents to send children to school, the question of whether and how tuition is free in the local compulsory primary education, the running of different types of schools, the strengthening of teacher training, the management of teachers in schools run by the local people, the collection of school funds for primary education, the construction of school houses, and the protection of the just rights and interests of schools. According to the local laws and regulations and in the light of the basic requirements of the Ministry of Education, the educational departments of the provinces, autonomous regions and municipalities put forward the specific requirements and methods for the investigation of the achievement of UPE in the respective areas.

2. Under the leadership of the peoples governments, the leading groups responsible for the investigation of the achievement UPE are set up at province, autonomous region, county and township levels with the participation of departments concerned with looking after the administration of the programme.

3. The leading groups in the provinces, autonomous regions and municipalities pick out one or two counties for trial investigation of the achievement of UPE, and are thus able to provide useful experience in the task of investigation on a larger scale and for various counties and villages.

4. The work of investigating the achievement of UPE is carried out from bottom to top. For primary schools in the cities and townships, the schoolmasters carry out a self-examination according to the basic requirements. Those schools which have reached the standards apply to the county educational bureau and then the examination is done by the district (or county). In the countryside, first of all the township leading groups check the primary schools in the local areas according to the basic requirements and the appropriate province, autonomous region or municipality regulations. Those schools which have reached the standards apply to the county educational bureau and then achievement is assessed by the county. For those schools which have reached the standards the counties would then apply to the provinces, autonomous regions or municipalities. The governments of the provinces, autonomous regions or municipalities would check and approve, or the work would be done by the respec-

tive educational bureaus which would then report to the next higher body for the record.

The provincial, autonomous regional, municipal and county leading groups for the examination of UPE generally go down to the schools, classes, villages and homes to investigate and check data during the examination process as well as hearing reports from the villages and schools. They make an inspection tour of the school yards, compare various tables, charts and names of the school rolls with the actual staff members and students to see if they match, and check the actual situation of those children who were unable to learn. They also visit classes, examine the teachers' teaching notes and the students' exercise-books, check and approve the test items and test papers for the graduating children, spot check the quality of education for those children outside the school, hold various discussions, visit teachers, pupils, parents, villagers and village cadres. Finally, the leading groups write out a report on the basis of the abundant information and figures then available.

The provinces, autonomous regions and municipalities directly under the central government, at the end of a year, hand into the Ministry of Education the list of counties where primary education has been popularized and all the relevant material.

5. The school, township and counties that have reached the standards for UPE are granted the quality certificates by the respective higher peoples' governments and for those schools and individuals that have made remarkable achievements. The various governments appraise and award them. The schools, townships and counties work out measures for improvement, enhancement and consolidation in the light of the shortcomings discovered during the investigation. Generally speaking, the governments of the provinces, autonomous regions and municipalities stipulate that for a period of three to five years an investigation will be conducted yearly in those schools, township and counties which have reached the standards. If the rates cannot be maintained in those places the quality certificates are withdrawn.

According to the standards set down, by October 1984, 552 districts attached to cities and 393 counties had reached the goal of UPE. According to the plans made by 28 provinces, autonomous regions (not including the Tibetan Autonomous Region) and municipalities, by the year of 1990, 2,179 cities and counties will have reached the goal of UPE which accounts for 91.8% of the total number of cities and counties.

III. Guiding Principles in the Establishment of the Basic Requirements for UPE

1. Strict demand and guarantee of quality

Like most of the developing countries, China has problems in the achievement of UPE. For example, educational funding is in short supply, the conditions for running schools are inadequate, there are not enough qualified teachers, and a considerable number of school children discontinue their studies. According to some figures, there are still 25,059,000 square metres of dilapidated classrooms. Quite a number of primary schools in China have not reached the standards laid down by the Chinese Government in 1980, that is, there should be no dilapidated classrooms, there should be a classroom for each class, and a table for each school child. Of 5.425 million primary teachers in China, qualified or semi-qualified teachers account for 48%, about 40% of the total primary teachers need training, and 12% of them are neither qualified nor deserve the retraining. The places of the latter teachers can only be taken by graduates from normal middle schools, and the readjustment will be completed only gradually.

In the whole of China, the enrolment rate of school age children has reached 95%, but the retention rate of children at school is only 93%. This figure means that in China an average of 11.65% of school age children do not attend school or discontinue their study. In the countryside, especially in remote mountainous, or backward areas the figure is even higher. In regard to the rate of successful completion of primary education, in the cities the figure is 90% but in the countryside it is only about 60%. For these reasons, in some places there are now more new illiterates than the former illiterates who have learned how to read or write. Therefore, when it worked out the basic requirements for UPE, the Ministry of Education stipulated not only the enrolment rate of school-age children and the retention rate of school children, but also the rate of primary school graduates and the rate of UPE among children between 12 to 15. It demanded that the schools actively take in illiterates or semi-illiterates of 12 to 15 of age back to school, or resort to various learning methods to help them to finish the major courses, or to recognize 2,500 commonly-used Chinese characters and be able to perform the four fundamental operation of arithmetics in round numbers in order to guarantee the quality of primary education and prevent the emergence of new illiterates. The recent work of investigating UPE in various provinces, autonomous regions and municipalities, show that it is rather difficult to reach the basic requirements in the countryside, especially in areas where economic, cultural and educational conditions are inadequate. However, with the effort of local governments, schools and teachers, the goal may be reached.

2. In view of China's situation, we should stick to the principle of combining uniformity with variety. Consequently, the provinces, autonomous regions and municipalities are allowed and required to set different requirements and to give appropriate guidance in the light of the unified basic requirements when they make plans for various counties. For the cities and rural areas where conditions are better, the requirements are set at a higher level and the goal of UPE is reached earlier. For

poor areas, the requirements may not be so strict and the goal of UPE may be reached at a later date. This principle was practised when the provinces, autonomous regions and municipalities worked out the specific standards for investigating the achievement of UPE. Especially in the cities and the areas where conditions are comparatively favourable, on top of basic requirements generally, the standards are set at a higher level. For instance in Beijing, after entering school children of school age are strictly not allowed to leave school before finishing primary education, except for special cases which are jointly approved by the township governments and district (county) educational bureaus. Of all the pupils at school, the rate of dropouts should be below 2%. In Jiangsu Province, in the regulations governing UPE, beside the rate of graduates, there is an additional rate of promotion to a higher grade in the non-graduating classes which should be 95%.

3. In the process of investigation, work had been done to create favourable conditions and to continually raise the level of primary education. Due to the shortage of teachers in large areas of the countryside except for mathematics and language courses there are no other courses stipulated by the Ministry of Education for the full-time primary schools. Such other courses could be natural history, history, geography, music, painting and physical education. The rate of graduates for the calculation of UPE among all the graduates is determined by pass or fail in the courses of mathematics and language. It is still a long-term task confronting China to continually raise the level of primary education and its universalization. Therefore, the Ministry of Education demanded that when investigation of the achievement of UPE was being conducted, effective measures must be taken to improve the conditions in the area of teacher training, funding, school houses and equipment, and the governments at various places must see to it that the work of UPE must be put on a solid and reliable basis. Those areas that have already attained universal primary education, must consolidate and improve what has already been achieved. Consequently, investigation of the achievement of UPE in the provinces, autonomous regions and municipalities and counties include not only the requirements for the four "rates", but also the following points:

(1) Have the governments at all levels put UPE on the agenda as an important task?

(2) Are the county (district) administrative and professional offices which are in charge of primary education functioning properly?

(3) Is there a rational layout of the township central primary school, six-grade primary and village school?

(4) Do the township central schools function as a model?

(5) Have the administrators such as the school masters and

others been provided, and are they competent? Is there a plan the teacher training and teaching research activities? Do the majority of teachers prove to equal to the job? Has the payment of the teachers working at school run by the local people been made?

(6) Is it true that there are no longer any dilapidated classrooms? Is there a classroom for each class, and is there a table for each pupil? Have the township central schools and the six-grade primary schools got necessary teaching equipment, apparatus, school walls and school gates?

The investigation of UPE in various places shows that due to the participation of the government leaders at all levels and the various departments in close connection with education, not only has the importance of UPE been understood and there is a common understanding of the matter, but also within a short period of time many problems that had not been solved for a long time have been easily solved. The plan for helping and training teachers, and measures to consolidate and develop what has been achieved are being worked out on the basis of the investigation. And all this has laid down a good foundation for further raising the level of primary education and its universalization.

IV. Problems that Need to be Tackled and Solved

1. The basic requirements stipulate that, upon graduation, pupils should be examined morally, intellectually and physically. The question of how to check moral and physical education remains to be solved. At present the standards for physical education testing are being made and are in trial use. Some time in the past, the general requirements were that the pupils should be in good health and have reached the standards in a few sports items like 50-metre running, long jump and hand-grenade throwing. In moral education, in most of the places pupils are asked to obey "Regulations for Primary Pupils".

2. In calculating the progress in UPE, the country is used as the basic unit. This helps to quicken the steps towards UPE in China as a whole and in the counties where the work of UPE has made great progress. But there are a few villages that have not reached the standards and have thus impeded the fulfilment of the task. However, in many counties where the goal of the four "rates" has not been reached, there are townships that have already reached the goal. These township are not included in the statistics and so the investigation will not be able to reflect the reality. Take 1984 for example, there were 945 city districts and counties that were checked and proved to be up to the standards, accounting for 36% of the total number of districts and counties. Yet in that same year, the enrolment rate of children of school age had reached 95%, the yearly retention rate over 93%, the rate of qualified primary school graduates over 90%

in cities and that of rural areas over 60%. Therefore, the question remaining to be answered was which level should be chosen as the basic unit for calculating statistics which will be beneficial to the promotion of universalization of primary education and a true reflection of the progress of UPE.

The Standard and Method of Assessing Pupil Achievement in Chinese Primary and Secondary Schools

Yang Dongliang

The assessment of pupil achievement is important in education. The assessment of pupil achievement at Chinese primary and secondary schools fall into three categories: behavioural record, academic record and sports record. These three aspects constitute a comprehensive index system of the assessment of pupil achievement.

The behavioural record refers to the examination and assessment of the pupil's moral character. A teacher comment is frequently used in China for assessing the pupil's moral conduct. The teacher in charge of a class examines his or her pupils' everyday moral conduct and writes a comment for each pupil at the end of a semester. The comments are based on "regulations for primary pupils" and "regulations for secondary pupils" laid down by the Ministry of Education. The comments generally cover three areas which:

1. affirm the pupil's merits and progress,
2. point out his or her shortcomings and deficiencies; and
3. make some suggestions.

In the early 1950s, the schools, apart from making comments, also gave grades from A to D, with C as a passing mark. In 1955, A five-grade system was used. In 1963, China issued "provisional working regulations regarding full-time primary schools" and "provisional working regulations regarding full-time secondary schools", which laid down that, in assessing pupils' moral conduct, only comments (not grades) should be used. The new method has been used until today. The experience of moral education of the past years has proved that comments can help parents to know their children better and help the students to give full play to their strong points, correct their own defects and thus encourage them to make constant progress. But there have been some shortcomings, for example, the comment is generally written by the teacher himself. As a result the comment tends to be subjective and naturally it is often one-sided. The pupils' moral records should include the record of moral conduct both at and outside school, especially a record of students' behaviour at home. The comment by the teacher cannot provide a full picture of the pupil's behaviour. In the past few years some Chinese primary and secondary schools have been doing some experiments, trying to make the assessment of pupil behaviour quantitative. For instance, the content of moral examinations is subdivided into several areas which are tabulated, with a numbering system for each area. According to the pupil's deeds, the pupil himself, his parent and teacher respectively work out a total mark and a grade, and the teacher finally writes out a comment. There are controversies over this new method. Those in favour say that there are marks in individual areas and a total mark which makes the comparison among pupils easy. Areas of deficiency can be easily identified. With the participation of the pupil himself, parent and teacher, there is self assessment and at the same time comments from the parent and teacher, so the comment is generally comprehensive. The teacher in charge of the class can understand his pupils better with information from all sides. Thus he is able to make comparatively objective and correct comments, hoping to make ideological work among pupils fit into the actual condition, in accordance with the pupils' characteristics. The ideological work thus seems more convincing. Some others argue that the scientific base of assessing moral conduct with marks can be doubted, for example, the questions of differential grades and use of marking system have not been satisfactorily solved.

The reform of pupils' behavioural record assessment method at present is conducted only on a small scale. In China as a whole, the assessment of pupils' moral conduct is in the main done by means of comments. It is very important to establish a scientific method and standards for assessing pupils' moral deeds. The nature of one's ideology and moral development involves a complex process and is determined by many factors, so the progress of pupils' ideological and moral development, unlike their knowledge, ability and physical condition, is not easily measured. All this makes it difficult to work out methods and standards of moral assessment. A set of scientific, rational and effective methods for assessing pupils' moral conduct in China does not yet exist. We have just begun research in this area and many of the problems need to be tackled in practice in order to be solved.

The assessment of pupils' achievement refers to the measurement of their level of knowledge and ability. The assessment of achievement of Chinese primary and secondary pupils includes quizzes at ordinary times (referring to classroom exercise, homework and periodic checks), mid-term examinations, end-of-term examinations and year-end examinations (for graduating pupils this means the graduation examination), and entrance examinations. The assessment falls into two types, that is, examination and quiz. In primary schools, for Chinese and mathematics, examinations are administered while quizzes are given in subjects such as ideological education, natural knowledge, history, geography, physical education, music, and drawing. In secondary schools examinations are administered in subjects such as political science, Chinese, mathematics, foreign languages, physics, chemistry, biology, history and geography. Physical education, music and drawing are examined in the form of quizzes. In every semester, for those subjects that have examinations, the achievement record follows this pattern: of the total, marks of quizzes at ordinary times account for 20%, mid-term examinations for 40% and the final examination for other 40%. In the year-end academic record, the record of the first semester contributes 40%, and the second semester 60%. The marking system is a 100-point system with 60 points as the pass level. For those subjects that require only quizzes, there are only term examinations and end-of-year examinations, and these subjects are not included in graduation or entrance examinations. The assessment of achievement is based on marks of weekly quizzes and of term examinations each of which is 50% of the total. The year-end academic record is calculated as the mean score of the two semesters. For most of these subjects, a 100-point system is used, while records for a few subjects are expressed in the form of grades. According to the teaching plan and syllabus, the schools assess the pupils' achievements. In the Chinese countryside graduation test papers are designed by the school district and the marking system is standardized. In the cities this is the responsibility of respective schools. In regard to the junior entrance examination, in the countryside the county educational department is responsible for designing test papers and marking system. In the cities the task is done by the district education department.

From what has been described above, it is clear that the assessment of achievement in Chinese primary and secondary schools is mainly done through examination. On the whole the requirement is strict and the standard is unified. However, there are problems:

1. As there are unified entrance examinations at all levels, there is often a tendency to attempt to enter a high proportion of students to schools of a higher grade. As a result, the development of students, morally, intellectually and physically, is neglected.

2. There are many types of tests and the tests are frequently given, thus putting much pressure on pupils.

3. There are many written tests and the testing method is not flexible. There are not enough adequate checks on the pupils' abilities. The defects in this area have already aroused the attention of the state and the public. In 1983, the Ministry of Education promulgated ten-point regulations regarding correcting the tendency of seeking to enter a high proportion of students to schools of a higher grade. It stressed that the education departments at all levels and the schools should adhere to the correct policy of running a school and take concrete and effective measures to guarantee an all-round development of students - morally, intellectually and physically. Some time before that, in some provinces, cities, municipalities and autonomous areas, excellent students had been excused from an examination and sent to senior high schools or universities to study, and various education departments also explicitly limit the times of

examinations at primary and secondary schools, stating that only two examinations can be administered in a semester, and that except entrance examinations, there should be no unified examinations. At the same time students' homework should be reduced and the amount of exercises should be controlled. It should be pointed out that reforms of testing systems are being conducted in some places and junior middle school entrance examinations have been abolished. The graduation examination has been merged with the junior middle school entrance examination. Many schools have started to reform the testing method in order to find out a better way of assessing pupils. The state is also actively involved in the research and improvement of test design, test content and marking systems, hoping to assess pupils' level of knowledge and ability in a more comprehensive and precise way.

In the teaching of Chinese primary and secondary schools, the utilization of examinations in assessing pupil achievement, and as a form of feedback, play important roles in raising the quality of teaching. However, many facts have proved that only by correctly using examination methods can it function effectively. Otherwise there are harmful results.

The assessment of physical culture refers to the examination of and comment on the condition of pupils' physiques and their sporting levels. For quite a long time in the past, the assessment of physical culture of Chinese primary and secondary pupils was based on their marks on physical education and their records of the state standards for physical training. The examination of primary and secondary pupils' physical education achievements includes a check on the knowledge of basic physical education and a check on the level of different sports items like running, jumping and throwing. In 1954, the state sports commission worked out and then promulgated "the state standards for physical training". Since then, part of the content of out-of-class sports at primary and secondary schools in China has been to put the standards into practice and the standards have become one of the important criteria in assessing the pupils' sports achievements. The pupils' records of physical education are combined with their record obtained in state standards for physical training to evaluate their physical culture achievement. This has given satisfactory results in helping the pupils to learn some basic knowledge of physical culture and sports techniques. But there are two major defects:

1. The evaluation is not comprehensive with a lack of the index to the pupils' physical condition. The fundamental task of physical culture at school is to improve the pupils' physical condition, so the check on the pupils' physical condition should be included in the evaluation of physical culture achievement.

2. It is not easy to get a correct proportion between the record of physical education and that of state physical training. The reason is that there are some overlapping sports items in physical education and the state physical training. Very often a

check on the standards replaces the check on physical education. According to some statistics, about 95% of Chinese primary and secondary pupils have reached the standards, which is much more than the state's requirement of 90%. Obviously, the present evaluation cannot adequately reflect the pupils' physical culture achievement. In order to set up new assessment methods and standards, since 1982, more than 40 primary and secondary schools in China have started some experiments according to "A provisional physical culture programme" designed by the physical culture department of the Ministry of Education. At the present this experiment has expanded into some 15 provinces. "A provisional physical culture programme" has in the main four points:

1. The record of physical education covers 50% of the total marks with 10% of basic physical culture knowledge and 40% of records of tested sports items.

2. The record of the state physical training covers 20%.

3. The physical condition, referring to physical shape and function (for example, height, weight, chest measurement, vital capacity and eyesight) covers 20%, with 5% for physical shape, 5% for physical function and 10% for eyesight. As to the standards for the check on physical shape and function we use index numbers rather than absolute figure. For instance, for physical shape we use weight/height x 1,000 for physical function, and chest measurement/height x 1,000. In checking eyesight we use an international visual chart.

4. The participation of out-of-class sports covers 10%. The "programme" has corrected the long-standing tendency to put too heavy a weight on the check on pupils' sports abilities at school, and the checking of the pupils' physiques has been added. As a result, it is easier to evaluate the pupils' physical culture achievement in a more scientific and reasonable way. Reports from those schools where experiments have been going on, show that the new method is superior over the old method of physical education record plus record in the state physical training. It helps to build up the pupils' health and raise their records in sports. From a long-term point of view, the content of the "programme" serves as an important reform in the work of evaluation of physical culture achievement at Chinese primary and secondary schools. The reform has gone on a correct path and the theoretical basis is also sound. But of course, the method is not without any shortcomings. For example, there are two many items to check, and the work needs the cooperation of physical education teachers, school doctors and the teacher in charge of a class. The load on the teachers is too heavy. At present, there is a universal shortage of primary and secondary physical education teachers and medical workers. There are considerable difficulties in putting the new evaluation method into practice. In order to solve the problem, it is imperative to improve the content of the marking system of the assessment method to try to make it easier to handle. At the same time physical culture

teachers and medical workers should be provided according to the establishment of respective schools as quickly as possible.

What has just been described is only a brief introduction to and simple analysis of the assessment methods and standards of the students' record at Chinese primary and secondary schools. Only a few more words are needed to state that the evaluation of moral conduct, student achievement and physical culture record forms an organic whole, none of them being dispensable. It is imperative to base the scientific assessment of student records on the index numbers of the above-mentioned three aspects and conduct a comprehensive and all-round assessment. We should not attend to one thing and lose sight of another, stressing only one aspect. The establishment of a scientific method of assessing students' records is an important task we face in the development of Chinese primary and secondary education. At present, work in this area has just begun and there are many problems that need to be tackled. We hope that by way of exchanging experience with other countries we shall learn more and carry this job of ours through to the end.

Educational Achievement in India

P.N. Dave

1. Background

India is a federal state and hence under the Constitution education is a responsibility of the State Governments. No sooner did the nation achieve independence, the founding fathers were seized with the problem of evolving a reasonably uniform system of school education, lest not only the standards of attainment by learners might vary but also there might emerge a citizenry with a narrow and regional outlook, thereby jeopardizing the hard-earned freedom, unity and national interests. The dilemma has been well articulated by the First Education Commission that was set up by the Central Government at the dawn of independence in October 1952. The Commission observed as follows:

> "...We recognize that Secondary Education is mainly the concern of the States but in view of its impact on the life of the country as a whole, both in the field of culture and technical efficiency, the Central Government cannot divest itself of the responsibility to improve its standards and to relate it intelligently to the larger problems of national life...Moreover, it is directly charged with the responsibility of maintaining proper standards in higher education. This cannot be done, unless careful consideration is given to the level of efficiency attained at the secondary stage." (The Secondary Education Commission--Mudaliar, 1952-53).

After a lapse of 12 years, the Central Government appointed a Second Commission, widely known as the Kothari Education Commission, to study the education scenario and recommend guiding principles and working policies for the development of Indian Education at all stages and in all respects. The Second Commission also found that the external examination dominated the implementation of curriculum. It voiced its concern of maintenance of standards and strongly recommended to define national standards at three levels--end of the primary stage (Class VII); end of the lower secondary stage (Class X); and end of the higher secondary stage (Class XII).

At this juncture India is in the process of enunciating a new education policy. In the document titled "Challenge of Education--a policy perspective" (1985), similar concerns about the variations of the standards from State to State and falling of and therefore improvement in the standards have been again raised. It has also drawn attention to the undesirable effect it has been producing not only on the implementation of curriculum primarily as a learning endeavour but also on the moral fabric of the national life. The failure of true examination reform, even after strong recommendations of the two national commissions and very active thrusts by NCERT at the school level and University Grants Commission (UGC) at the university level indicates that the Indian Educational System has developed resistance to any change. It is not the dearth of ideas or availability of viable strategies that are impending its reform. But it is the lack of political will and the intent to implement the well conceived schemes of evaluation which has created an impasse. Vested interests, particularly on the part of the teaching community at all levels of education, have not played a small role in this state of affair.

2. Present Status

2.1 Assessment at the Primary Stage

A mention has been made above that India is a federal republic and, according to the Constitution, education is the responsibility of the State Governments. The data regarding the evaluation practices at the primary stage is scanty. However, except in a few States and Union Territories, at least two examinations are conducted at the end of two sessions of six months, generally known as the half-yearly or terminal examination and the annual examination. Some are continuing the past practice of conducting a common examination at the end of class VII at the district level.

Internal Assessment: The internal assessment of children is conducted periodically. But there is a wide variation in this regard. The good schools, particularly in the urban areas, carry out this task very systematically, while the poor schools, especially in the rural areas, fail to do so. This results into creating uneven standards of achievement of children.

2.2 Assessment at the Secondary Stage

Public examinations are conducted at the national and state levels at the end of class X and XII in most of the States and Union Territories. In a few States, which have not yet switched over to the 10+2 system of school education, these are conducted at the end of class XI. Like in the primary classes, two terminal exams at the end of each six-month term are conducted. The internal assessment comprises of teacher-made tests periodically administered to the students. The marks/grades of these tests are given some weightage in declaring the final results.

Purpose: By and large, these external and internal examinations are found to be a routine affair. The focus appears to be on declaring the learner as "passed" or "failed". These fail to achieve the most important purpose of ascertaining the extent to which the desired educational objectives have been achieved. Rarely are they used to identify the strength and weakness of the learner. On the contrary, the word examination, may it be an external, internal or a test, arouses fear and anxiety in the mind of the learner, thereby, more often than not, disrupting the normal response system of the learner during the examination. It is probable that it may result into developing undesired work habits, attitudes and values in the learners.

Emphasis on Scholastic Achievement: There is undue emphasis on testing mostly the growth and development in the cognitive (scholastic) area during the years in the school. By and large, the development in the Psycho-motor and Affective Domains are not given sufficient recognition. Consequently, the development and assessment of skills, particularly manipulative ones, like using tools, apparatus, equipment, etc. and interest, attitudes and values like initiative, cooperativeness; industriousness, spirit of social service, social justice etc. are not properly attended to in the school. No due weightage is given to the achievement of children in the non-scholastic area in the certificate issued to the children at the end of each year or at the common public examination at the end of school education. A typical certificate carries marks, or grades in a very few cases, only in the scholastic subjects studied. The marks in the practicals and work experience/Socially Useful Productive Work are shown, as they form the part of compulsory subjects, but are neither taken into consideration for admission to the higher classes or colleges nor given much importance by the employer while selecting the candidate for a job. Even the high attainment in languages, so useful a communication tool, is ignored when admission to non-humanistic areas of learning is sought by the learners. Thus, the entire school education--the transaction of the intended curriculum and its evaluation--is confined to teaching and assessing a few scholastic subjects, leaving thereby the development of many a desired socio-emotional and physical characteristics/traits lopsided.

3. Review of Related Literature

Surveying the Indian educational scenario dispassionately, one finds quite paradoxical situations existing side by side. In contrast to the gloomy picture of the implementation of the policies, the picture of theoretical and empirical research in the area of examination and evaluation, although based on the Western concepts and strategies, is much brighter and more impressive. Due to the serious efforts of documenting the research in Education and allied fields and periodical upgrading of the same by NCERT (1966), Pareek and Kumar (1966), Pareek and Sood (1971), Dave, R.H. (1968), Buch (1972, 1979 and 1985), substantial and significant findings in respect of various areas of education are readily available to the research worker.

The Third Survey of Research in Education (Buch, 1985) brings forth up-to-date information on the subject under the chapter entitled "Educational Evaluation and Examinations", authored by Passi and Hooda. The data presented in Table 1 show the distribution of 198 researchers carried out between 1949 to 1983 under the six areas, viz., Achievement Test, Diagnostic Test, Examinations, Factors Affecting Achievement, Prediction-admission-promotion and Failures. Mention needs to be made that the studies comprise mainly Ph.D. these submitted to various Indian Universities and independent studies undertaken by institutions and individual scholars.

Table 1: Areawise and Periodwise Distribution of Studies

	Upto 1949	1950-1954	1955-1959	1960-1964	1965-1969	1970-1974	1975-1979	1980-1983	Total
Achievement Test			2	10	16	17	16	8	69
Diagnostic Test			-	-	5	5	2	2	14
Examinations			2	11	12	13	15	9	62
Factors Affecting Achievement	1	2	5	2	4	3	2	1	20
Prediction-Admission-Promotion	-	2	2	6	4	3	3	1	21
Failures		-	2	4	2	-	3	1	12
Total	1	4	13	33	43	41	41	22	198

In the First Survey, Buch and Yadav (1974) provided some interesting information. While the Ph.D. studies conducted until 1972 comprised only 9.6% of the total number of 342 Ph.Ds, those carried out as independent research constituted 22.3%, the highest of the total number of 269 studies. This strongly indicates the attention given and the importance attached to this area by the research workers. Passi and Padma (In Buch, 1974) reported that 95 studies on the subject were taken up between 1944-1972. The cumulative figure of studies carried out between 1944-78 was 127, showing an increase of 32 studies (Passi and Sansanwal, In Buch, 1978). Passi & Padma (In Buch, 1985) gave the cumulative figure of studies carried out between 1944-1983 as

198, indicating thereby an increase of 71 researches during the gap of five years. One can conclude from this information that this area is receiving quite adequate attention both from young researchers and eminent scholars.

The data in Table 1 shows that within this area, examinations (62) have received as much attention as has been by achievement tests (68). Thus this sub-area also has been receiving more than adequate attention.

The first study in the area was conducted by Chickermane in 1943 (Passi & Padma, 1972). Since then a few individuals and institutions have made sustained contributions to this field. To name a few, Bokil at the Maharashtra State Board of Secondary Education (1956-1963), Buch and Lele at the M.S. University, Baroda (1960-1972), Harper at the Ewing Christian College Allahabad (1960-1970), Gayen at the IIT, Kharagpur (1961-1970) and Taylor at Guwahati University (1963-64), did commendable research work. Most of all the studies carried out by these researchers were supported by NCERT or University Grants Commission. NCERT on its own has made a valuable contribution to R and D efforts in this area at the national level.

Passi and Hooda (1984) have analyzed the findings of all 198 research studies and discussed the trends with regard to the following characteristics: (i) Validity of the measures, (ii) Reliability of the measures, (iii) Effect(s) on curriculum and teaching and (iv) Comparability of standards.

Validity and Reliability of Examinations

The analysis of results indicates that the validity and reliability of the present examination system is very low, as the stress has been on the performance of the learners in a single-shot-final examinations.

Narrow Coverage of Objectives: Even in the assessment of scholastic achievement, not full justice is done to the learners, for, only a limited number of objectives are tested. The design of question papers, until recently dominated by the essay type questions, is such that mostly rote memory or lower mental abilities confined to the storage and retrieval of information ('knowledge' in the Bloomian terminology), are tested. A very little testing, if at all, of understanding, application and creativity is even attempted in the external examination. On a very small scale and proportion, that too in a selected progressive schools, attempts are made to expose children to testing situations which cover hierarchical levels of mental abilities. The Indian school system functions under severe resource constraint, and therefore has a very unequal distribution of threshold educational facilities--physical and academic. This results into wide variations of curriculum transaction which in turn create uneven levels of achievement.

Narrow Coverage of Content: Yet another lacuna in the assessment of the learners, particularly in the external examination, is limited coverage of the content. The one-shot examination severely suffers from the constraint of time. A three-hour question paper cannot cover all the essential concepts, principles, processes and critical content relating to subjects and hence the paper setters resort to selection, not infrequently using subjective criteria. Besides, even in the inadequately designed question paper, options are (perhaps have to be) provided, which reduces the content covered by introducing an element of further selectivity from the already reduced content coverage. This is further accentuated by other factors such as scoring by a large number of different examiners, resulting into clustering of marks at the critical 'pass' of 'fail' line, usually 33% marks, followed by big gaps below the same. Ultimately what is being tested is a very small fragment of scholastic subjects. This is indeed, a very serious flaw that affects the assessment of the cognitive growth and development which has been the prime focus of the educational system.

Inter and Intra Examiners Reliability: Time and again it has been demonstrated that the fully essay type of examinations are not reliable, whereas the objective type examinations are more reliable and valid. Inter examiners reliability has been found very low. One researcher made a very strong comment stating that "an examiner's marks has neither the sanctity nor the precision which is usually attached to it". The means and S.Ds of different evaluations of different examiners widely vary. However, most of the studies indicate that the consistency among assessments by the same examiner is fairly high, even with the essay type examinations. It increases if grades are assigned instead of marks.

Design of Question Papers: In the early sixties, NCERT heralded the movement of examination reform by establishing a separate department of evaluation. With the help of Benjamin Bloom, a series of seminar-cum-workshops were organized to train the high level state functionaries, particularly those manning the Board of Examinations. It was envisaged that within two decades, the examination system would rid of its weaknesses and a systematic, continuous and comprehensive internal evaluation system will supplement the external system, if not replace it.

The long experience of the last two decades has belied those expectations. While a substantial number of officials have been trained all over the country, the impact of the movement within the system is minuscule.

General observations have been made about the design of question paper(s). Several studies show that a large number of failure can be attributed to the defective question papers and careless or even incompetent paper setters. In general, it has been reported that the question papers lack balance, discriminating power, adequate coverage and appropriate weightages. Even alternative forms of question papers set for the forenoon and afternoon sessions are not analogous and of the same standard.

Thus a good theory and its operational steps for designing objective-based question papers are available; but their implementation has not been possible. The change in the overall format with due weightages to behavioural objectives and content of a subject, and the increase in the total number of questions and also the types of questions within a paper, have been met with stiff resistance. These have also created problems of management, cheating, mass copying, corruption etc. in the organization of public examination as well as the internal assessment to a smaller scale. For example, objective type of questions such as multiple-choice, true-false etc. have been abandoned because of copying. Paradoxically, the same have been effectively used for the entrance to the engineering and medical colleges/institutes and the selection for Indian Administrative Services (IAS), Defense Services, Bank Services, etc. As a matter of fact, as culmination of such developments in the field, an idea of National Educational Testing Services has been emerging as a viable alternative to the present system of public examination. This would certainly help in improving the design and format of both questions and question papers.

Comparison of Awards of Classes/Marks/Grades: The problem of variations in the level of achievement among the subjects under an examination, among the same examinations from time to time, among the examinations conducted by several Boards of Examinations and evolving a mechanism for developing and maintaining the national standard, has been receiving serious attention since the Constitution of the First Secondary Commission after Independence.

The problem in entirety is vividly exposed by Arora (1985) in a statistical presentation of the comparative data of passed students in examinations conducted by 25 State Boards of Examinations during 1980 and 1981 (see Table 2). A quick scrutiny of these results shows that: (i) The overall percentages of passed students in 1980 and 1981 are only 48.63 and 48.18, whereas the same for 1978 and 1979 are 47.15 and 37.15 and 37.43 respectively, (ii) The range of pass percentages varies from 39.98 (Assam) to 91.45 (Council for the Indian School Certificate Examinations, Hill Schools only) in 1980 and from 16.71 (Madhya Pradesh) to 93.92 (Central Board of Secondary Education, New Delhi) in 1981 respectively.

Other researchers have reported variations among pass percentages over a five-year period from 1960, increase in the average performance in Hindi (language), mathematics and science and significant variations in the percentage of passes among subjects over years. However, further examination of the data indicated that the important reason has been the nature and difficulty levels of the question papers and not perhaps the increase in attainment as such on the part of the learners.

It is not out of place to highlight some stark facts regarding the wastage and stagnation of human resources that emerge from these statistics. India is struggling to achieve universalization of elementary education, i.e., classes I to VII. While around 83 per cent of the eligible children enrol in class I. Only about 40 per cent of these children finish class V, and about 22 per cent of the same cohort group manage to complete class VIII. However, about 83 percent of these children carry on their studies upto class X. The most generous estimate of children taking this examination would be around 15 percent of those who should have entered class I ten years ago. One would not hesitate to label this group as bright children. Now the result of Boards of Examinations have been consistently around 50%. Even if the entire eligible age group would have reached class X, one would have expected more than 70% of the children to pass the examination on the statistical ground, i.e. the assumption of normal distribution. The failure of the half of the academically quite successful children reflects very adversely not only on the examination system but also on the whole education system. It is indeed a tremendous waste of human resources.

Internal and External Assessment: A fair number of studies have been carried out in order to find out the relationship between them both at the school and college levels, in different faculties, in different geographical areas and with various types of examinations. Mean differences, pass-fail differences and correlations have been used to analyze the results. The trends that have emerged are: (i) there has been liberal making in the internal assessment, (ii) the relationship between these varies from institution to institution subject to subject and teacher to teacher and (iii) the internal marks possess some predictive value.

Theory and Practical Examinations: A small number of studies have been conducted to investigate the relationship between the marks obtained in the theory and practical examinations. These studies reveal that there is a wide disparity in making in these two examinations. Like in the internal assessment the marking in the practical examinations has been found quite liberal. However the coefficient of correlations have not been very high.

Failure: A reference has been made somewhere about the disproportionate number of failures at the national and State level examinations, resulting into an enormous wastage of resources. This phenomenon has attracted the attention of research workers. In general, the findings indicate that failure in the examination is the consequence of the interplay of a large of number variables related to pupils, teachers, subjects, schools, parental status and the examination system. Some specific variables need mention here: size of school, subjects of English and Mathematics in general, socio-economic status and inadequacy of mastery over the subject by the teacher.

Mass Copying: Mass copying along with devious means of cheating has been emerging as a hot issue. It is but natural that attention of the researcher has been drawn to it. A lone investigation has revealed that those who are involved in mass copying come from homes of low socio-economic status, low parental education and having inadequate facilities.

Innovations: As may be expected, having done intensive research work in the field, some reputed researchers have made concrete suggestions to improve the traditional examinations. The scaling of marks was recommended by Taylor; screening of examiners by Bose; appointment of independent examiner for each question by Harper; and use of standard error by Taylor and Tluanga. A suggestion of the vertical system of promotion of an individual pupil, stage by stage and subject by subject has also been mooted by others. The acceptability of 13 innovations was studied at the university stage. The following have found favour with both students and teachers: Introduction of the grade system of ranking, supplementing essay type with objective type examinations, semester system, internal assessment as supplementary to external examinations and defining the scope of questions in simple and clear language.

Question Banks: The major weaknesses of the examination system as evidenced by researches have been discussed above. Recently, a new development has been taking place. Various institutions and individuals have been attempting to develop question banks at the school and college levels. It is just the beginning and the data regarding their usefulness, reliability and validity have to be systematically collected before any judgement is pronounced for or against them. However, it is necessary to record here that, although UGC mooted the idea long ago, hardly any mass scale reforms are attempted at the university level. It is not an exaggeration to observe that perhaps the university sector or higher education sector appears to be the most conservative and therefore least amenable to change.

Table 2: Number of Candidates Passed in 1980and 1981 Examination at the End of Class X in Boards of Secondary and Higher Secondary Education

Name of Board	Passed	
	1980	1981
1. Office of the commissioner for Govt. Examinations, A.P., Hyderabad	148,295 (50.64)	186,962 (55.00)
2. Board of Secondary Education, Assam	65,291 (39.98) (Examination 1980 held in 1981)	N.A --
3. Bihar School Examination Board, Patna	--	135,957 (57.16)
4. Central Board of Secondary Education New Delhi		
i) All India	32,459 (88.16)	37,951 (93.92)
ii) Delhi Secondary	37,699 (55.86)	40,894 (56.44)

Table 2: Number of Candidates Passed in 1980and 1981 Examination at the End of Class X in Boards of Secondary and Higher Secondary Education (continued)

Name of Board	Passed	
	1980	1981
5. Council for the Indian School Certificate Examinations i) Plains School only ii) Hill schools only	 11,894 (87.08) 1,223 (91.54)	 11,451 (91.18) 1,360 (91.03)
6. Goa, Daman & Diu Board of Secondary & Higher Secondary Education, Goa .	4,169 (57.72)	4,625 (60.82)
7. Gujarat Secondary Education Board, Gandhinagar	120,760 (52.51)	118,988 (49.57)
8. Board of School Education, Haryana	45,237 (52.72)	63,437 (61.43)
9. Himachal Board of School Education, Simla	16,040 (56.13)	18,625 (63.62)
10. Board of School Education, Jammu & Kashmir	11,599 (36.43)	13,376 (40.97)
11. K.S.E.E. Board, Karnataka, Bangalore	139,936 (53.33)	68,832 (36.68)
12. Board of Public Examinations, Kerala	162,750 (42.58)	184,441 (39.49)
13. Board of Secondary Education, M.P., Bhopal	224 (44.53)	1,040 (16.71)
14. Maharashtra State Board of Secondary & Higher Secondary Education, Pune	212,873 (47.13)	247,780 (50.37)
15. Board of Secondary Education, Manipur	11,120 (56.43)	8,419 (44.61)
16. Meghalaya Board of School Education, Tura	2,083 (35.88)	2,639 (41.87)
17. Mizoram Board of School Education, Aizawal	1,426 (41.39)	1,901 (49.53)
18. Nagaland Board of School Education, Kohima	1,173 (40.74)	1,542 (42.89)
19. Board of Secondary Education, Orissa	54,025 (43.15)	72,819 (48.29)
20. Punjab School Education Board, SAS Nagar, Mohali, Distt. Ropar	8,994 (51.64)	99,344 (55.36)
21. Board of Secondary Education, Rajasthan, Ajmer	70,035 (42.20)	81,776 (42.67)
22. Board of Secondary Education, Tamil Nadu, Madras	136,276 (57.60)	150,575 (58.00)
23. Tripura Board of Secondary Education, Agartala	5,166 (48.68)	7,199 (50.91)
24. Board of High School and Intermediate Education, (Uttar Pradesh)	385,988 (45.15)	304,511 (38.38)
25. West Bengal Board of Secondary Education, Calcutta	130,742 (52.90)	127,659 (51.69)
	1,890,477 (48.63)	1,994,103 (48.18)
	1978	1979
TOTAL	1,677,778 (47.15)	1,828,338 (37.43)

(Figures within brackets indicate percentages)

4. National Schemes of Importance

After having studied the innumerable findings surfacing from the Ph.D these and other independent studies, one is left with a feeling that while significant bits of valuable information are emerging, the holistic pattern of configuration binding or interrelating these parts is not clearly perceptible. The Commissions have perhaps, without much empirical data, comprehended the problem in totality, which somehow seems to be missing in the way the researchers are conducted and/or the findings are presented and the trend reports on the subjects are written. It is a sad commentry on the system which can not provide solutions to the problems relevant to the national concerns. It is of paramount importance that a nation of wide diversities and disparities should have answers to the questions like--what is the achievement of an Indian child who has completed class I, or the lower primary stage or the primary stage in various school subjects? What is the attainment of graduates in different disciplines at the university level?

Except Shukla's lone study (1974) no serious attempt has been made to study the problem of achievement and standards at the school stage. Incidentally, the investigation conducted by Shukla (1974) was the part of a study sponsored by the International Association for the Evaluation of Educational Achievement (IEA) in their 1970 Science Survey (Comber and Keeves, 1973). Even this study was not addressed to all the states, as it was confined to the Hindi speaking region, comprising Uttar Pradesh, Bihar, Madhya Pradesh, Rajasthan, Haryana and the Union Territory of Delhi.

Unfortunately, such a valuable exercise in the area of educational achievement did not get sustenance in the country. Had it been institutionalized, valuable data for comparability of standards and their subsequent improvement and projection would have been possible. The data of the study in general showed that the achievements of Indian children in Hindi (Language) and Science at both the stages -- primary and secondary -- were significantly lower than their counterparts in the developed countries.

Achievement at the Primary Stage

The last decade has seen the implementation of massive innovative UNICEF assisted projects at the primary stage, covering virtually the entire country. "Primary Education Curriculum Renewal" is one such school project which was initiated in 1975-76. The project aims at developing innovative curricula and instructional materials relevant to the needs of diverse groups of children. Efforts are also made to adjust the existing curriculum qualitatively to the life-style of the child and socio-economic opportunities likely to be available.

For the last six years the project has been implemented in 22 States and eight Union Territories with the help of State Council of Educational Research and Training (SCERTs)/State Institutes of Education. Project activities such as socio-economic and educational surveys, identification of local specific content/situations, preparation of instructional materials, tryout of the materials, teacher training etc. are carried out involving 180 teacher training institutions and 2469 primary schools located mostly in the distant, backward areas of the country. The curricula have been developed on the basis of a specially and collectively developed document entitled "Minimum Learning Continuum" containing a list of graded and sequential competencies to be attained at the mastery level.

As the project progressed over years UNICEF, the Ministry of Education and NCERT collectively felt the need for evaluating the impact of the project, for, further steps of infusing the project ideas into the regular primary education system could not be taken in the absence of valid evidence of its effectiveness. Hence, a massive evaluation study was planned and organized in 1984 to ascertain the status of enrolment, retention, stagnation and achievement of pupils in the participating schools. In order to determine the effectiveness of the implementation of the renewed curriculum in the project vis-a-vis the non-project schools, a smaller study-design was carved out of the major study. The sample of two project and two non-project schools from the same homogeneous cluster within each state, was carefully selected at random. The children studying in classes I to IV in the schools of 23 States and Union Territories were administered a battery of competency-based tests in language, mathematics and environmental studies (Science and Social Sciences) under a very careful supervision of staff members from NCERT, SCERT, local teacher educators and teachers. The scoring and tabulation of data have been done at the state level, while the statistical analysis and computer processing are being carried out at NCERT.

Major Results: The data pertaining to pupil achievement in Language, Mathematics and Environmental Studies (Science and Social Sciences) are presented in Table 3, 4 and 5 respectively.

Achievement in Language: The range of percentage means in respect of classes I to III in different States/UTs (see Table 3) indicate that the levels of achievement in language vary widely. However, the combined percentage means of pupils in classes I and II (64.20% and 56.13%) suggest that the overall level of attainment is quite good. The data further indicates that the attainment level of pupils progressively goes down in classes III and IV, barely touching so famous a critical line of passing marks of 33%(32.50). Further, the children in the project schools have done significantly better in classes II & III than their counterparts, indicating thereby the effectiveness of the intervention of the project.

Table 3: Pupil Achievement in Language in Classes I to IV of Project and Non-Project Schools of Participating States & Union Territories

CLASS	Type of schools	Number of Students	Combined Mean (%)	Combined SD	Range of State Means	't' Value*
I	Project	1925	64.50	26.38	32.58-83.80 (Bihar)	.710 df= 58 n.s.
	Non Project	1814	63.89	25.13	41.22-79.58	
	Total	3739	64.20	25.79	19 States/UTs	
II	Project	1681	58.21	24.73	20.31-83.29 (Bihar)	3.244 df= 50 Significant at .01 level
	Non Project	1581	53.92	23.19	29.24-76.44	
	Total	3262	56.13	24.09	15 States/UTs	
III	Project	1346	42.94	22.33	17.36-69.60 (Bihar)	2.304 df= 37 Significant at .05 level
	Non Project	1325	40.92	20.66	20.35-62.70	
	Total	2671	41.94	21.54	12 States/UTs	
IV	Project	78	35.83	21.44	-	
	Non Project	127	29.16	14.59	-	
	Total	205	32.50	18.02		

* The school mean was taken as a unit for computation.

Table 4: Pupil Achievement in Mathematics in Classes I to IV of Project and Non-Project Schools of Participating States & Union Territories

CLASS	Type of schools	Number of students	Combined Mean (%)	Combined SD	Range of State Means	't' Value*
I	Project	1943	63.83	24.57	34.61-80.68	1.993 df= 58 n.s.
	Non Project	1806	58.44	26.47	31.02-77.61	
	Total	3749	61.23	25.65	19 States/UTs	
II	Project	1668	70.42	19.33	40.89-93.15	2.818 df= 50 Significant at .01 level
	Non Project	1-91	66.88	20.45	45.87-83.36	
	Total	3259	68.69	19.96	15 States/UTs	

Table 4: Pupil Achievement in Mathematics in Classes I to IV of Project and Non-Project Schools of Participating States & Union Territories (continued)

CLASS	Type of schools	Number of students	Combined Mean (%)	Combined SD	Range of State Means	't' Value*
III	Project	1355	51.56	27.38	15.23-75.85	1.143 df=37 n.s.
	Non Project	1348	50.96	25.97	33.95-77.99	
	Total	2703	51.26	26.69	12 States/UTs	
IV	Project	138	39.22	24.99	32.18-48.38	
	Non Project	209	27.97	16.60	24.68-33.07	
	Total	347	32.45	21.08	2 States/UTs	

* The school mean was taken as a unit for computation

Table 5: Pupil Achievement in Environmental Studies in Classes I to VI of Project and Non-Project Schools of Participating States and Union Territories

Subject	Class	Type of school	Number of students	Combined Means (%)	Combined SD	Range of State Means	't' Values*
EVS	I	Project	1949	76.86	23.03	51.32-93.95	3.076 df= 57 Significant at 0.1 level
		Non Project	1792	71.82	22.81	53.77-88.28	
		Total	3741	74.48	23.06	19 States/UTs	
	II	Project	1681	74.57	17.67	47.86-89.29	4.884 df= 50 Significant at 0.1 level
		Non Project	1553	65.18	21.74	42.46-87.85	
		Total	3234	70.06	20.28	15 States/UTs	
EVS-I (Social Science	III	Project	1342	52.58	26.77	18.54-81.49	1.834 df=37 n.s.
		Non Project	1321	49.42	27.74	19.66-84.12	
		Total	2663	51.01	27.40	12 States/UTs	
EVS-II (Science)	III	Project	1338	48.49	25.57	16.58-76.65	2.16721 df= 37 Significant at 0.5 level
		Non Project	1321	47.70	23.90	28.90-74.28	
		Total	2659	48.10	24.76	12 States/UTs	

Table 5: Pupil Achievement in Environmental Studies in Classes I to VI of Project and Non-Project Schools of Participating States and Union Territories (continued)

Subject	Class	Type of school	Number of students	Combined Means (%)	Combined SD	Range of State Means	't' Values*
EVS-I (Social Science)	IV	Project	138	35.77	19.73	33.95-38.13	
		Non Project	206	29.27	14.89	24.22-37.39	
		Total	344	31.88	17.29	2 States/UTs	
EVS-II (Science)	IV	Project	138	36.49	23.60	28.00-47.53	
		Non Project	208	28.64	16.05	21.40-39.98	
		Total	346	31.77	19.79	2 States/UTs	

* The school mean was taken as a unit for computation.

Achievement in Mathematics: The data presented in Table 4 reinforces the trends observed in respect of the data pertaining to the achievement in Language. However, the achievement of pupils in mathematics in class II is better than that in class I. There is not a progressive decline in the achievement, like in Language. All the same, there is a sudden slump in achievement in Mathematics in class III, again hitting miraculously the critical line of passing marks of 33% (32.45). Here also, the children exposed to the renewed relevant curricula and suitable strategies of teaching them have done significantly better in class III than those who were not.

Achievement in Environmental Studies (Science and Social Sciences): The data presented in Table 5 further confirms the trends of pupil achievement observed in respect of Language and Mathematics. The achievements of pupils in classes I & II are quite high. Here again the children in classes I & II studying in project schools have attained significantly higher percentage means than those studying in the non-project schools.

The overall results strongly suggest that, while the national achievements in Language, Mathematics and Environmental Studies in classes I & II are quite satisfactory or even touching the mastery in the subjects, the indices of achievements are extremely poor in the States like Bihar and Rajasthan, ranging from 15% to 35% only. Needless to mention, intensive efforts are required to raise the level of attainment in these States which are also bracketed as "educationally backward States" in view of low enrolment ratios and high dropouts rates.

The baffling trend has been the sudden drop in the achievement in class III in all states and in all subjects. It is irrespective of the type of schools. This necessitates an indepth analysis for ascertaining the underlying causes.

National Talent Search Scheme (NTS)

It might seem strange that after having lamented upon a lack of political and academic will and the impediments coming in the way to reform our monolithic public examination system, the facts, figures and findings of a truly national scheme are presented at the end. The very reason that it is outstanding and deviant from a routing, normal organization of examination, it required a special presentation in a paper of this nature.

Origin: The National Science Talent Search Scheme (NSTS) was launched in 1963 with a view to "identifying a team of brilliant students at the end of school stage of education who could be nurtured as the future quality scientists of India" (Saxena, 1967). This scheme was later renamed as National Talent Search Scheme (NTS) in 1977 in order to widen the scope of identifying talent in Social Sciences. The first scheme awarded the scholarships on the basis of the performance of students on the following components, viz., (i) A Science Aptitude, Test, (ii) An Essay Paper, (iii) A Project Report, and (iv) An Interview. The subsequent NTS comprised the following components: Part I, General Mental Ability Test (GMAT) of 100 marks, Part II, Scholastic Aptitude Test (SAT) of 100 marks and Part III, An Interview of selected candidates of 25 marks.

Extension of Scheme: In 1979, another significant change was introduced by extending the award of scholarships for the pursuit of higher education in professional courses such as Medicine and Engineering. Till 1982, this modified scheme continued as such except for increase in the number of scholarships from 550 to 750 in the year 1983.

The organization of the scheme until now was done centrally. As per the recommendations of a recent Review Committee, the scope of the scheme has been extended to all existing courses of study except Fine Arts in addition to Science and Social Sciences. However, its operations will be decentralized to the State/UT level.

Salient Features: The scheme has made significant contributions to the development of comprehensive achievement and aptitude tests. These are scientifically designed apportioning due weightages to higher objectives and content. The coverage of objectives is extensive and that of content broad. Thus the internal consistency of the tests has been enhanced. The variation in scoring is reduced to minimum except due to chance, as only the objective type tests, consisting of mostly multiple-choice items are used. Further the scope of the cognitive area is expanded by using the General Mental Ability Test. The assessment done during the interview has also helped in enlarging the scope as well as partially checking the reliability of the written tests. However, it may be noted that a good measure of other abilities such as creativity in the form of "A Project Report" had to be discontinued in view of the malpractices, although the reason given for dropping it was "to reduce subjectivity in assessment." In a nutshell, the conduct of the scheme has demonstrated the potential of implementing a good theory of evaluation. It has gathered valuable experience in the development of the design, execution and management of a massive national scheme that could become a basis for establishing a National Educational Testing Service.

Standards of Achievement: This is not the place of presenting details regarding the validity and reliability of the scheme. Such details are available in the collection of studies conducted by Misra et al. (1981). Two major points relevant to this seminar need to be highlighted. The data on predictive validity of the tests are not very encouraging. The follow-up research of selected scholars stared rather late. When the post-facto studies were initiated, a number of files of candidates could not be located. The most disturbing fact reported (Raina & Mittal, In Misra, 1981) is that out of 3988 candidates selected between 1967 and 1976, 2672 (67%) dropped out of the scheme. Although this phenomenon can not be attributed to the testing scheme, only it has partially contributed to it. However unfair it may sound the common man will find fault mostly with the selection procedure. A question comes to mind: Is it because the non-scholastic factors are not given their proper place in the selection of talent?

The second point relates to the standards of attainment by different States/UTs. The distribution of selected candidates in 1985 according to States/UTs is given in Table 6. Before the introduction of the new decentralized scheme, the Union Territory of Delhi used to bag the largest number of scholarships. For example, it secured 117 out of 340 general scholarships for class X, 57 out of 136 for class XI and 77 out of 204 for class XII during 1985. This terribly lopsided capture of scholarships by the Delhi candidates may have been one of the reasons for decentralizing the scheme with apportioning the quota on the basis of the number of enrolled/appeared candidates at the secondary school examination. The data in Table 6 once again confirms the well known generalization that the standards differ from State/UT to State/UT and the candidates from many States/UTs are quite below the standard expected to be eligible for the national scholarship. For illustration, out of 55 eligible candidates from Delhi 50 have been selected. In contrast to this, eight States/UTs have no candidates qualifying for an interview call. On the other hand, the State of Maharashtra has shown commendable improvement, thereby benefiting from the new scheme by slicing a bigger share of the talent cake.

Table 6: Statement showing number of candidates selected under the decentralized National Talent Search Scheme 1958

State/UT		Quota allotted	Selected			Percentage
			General	SC/ST	Total	
1.	Andhra Pradesh	190	33	06	39	20.5%
2.	Assam	130	04	-	04	3.1%
3.	Bihar	155	54	14	68	43.9%
4.	Gujarat	165	06	01	07	4.2%
5.	Haryana	60	18	-	18	30.0%
6.	Himachal Pradesh	25	04	-	04	16.0%
7.	Jammu & Kashmir	25	-	-	-	0.0%
8.	Karnataka	140	46	04	50	35.7%
9.	Kerala	275	49	04	53	19.2%
10.	Madhya Pradesh	225	26	03	29	12.9%
11.	Maharashtra	305	148	24	172	56.4%
12.	Manipur	25	-	-	-	0.0%
13.	Meghalaya	25	01	-	01	4.0%
14.	Nagaland	25	-	-	-	0.0%
15.	Orissa	110	26	02	28	25.5%
16.	Punjab	100	23	01	24	24.0%
17.	Rajasthan	150	39	05	44	29.3%
18.	Sikkim	25	02	-	02	8.0%
19.	Tamil Nadu	235	68	03	71	30.2%
20.	Tripura	25	05	01	06	24.0%
21.	Uttar Pradesh	445	30	01	31	6.9%
22.	West Bengal	175	39	01	40	22.8%
23.	A.N. Islands	10	-	-	-	0.0%
24.	Arunachal Pradesh	10	-	-	-	0.0%
25.	Chandigarh	10	04	-	04	40.0%
26.	D & N Haveli	10	-	-	-	0.0%
27.	Delhi	55	50	-	50	90.9%
28.	Goa, Daman & Diu	15	02	-	02	13.3%
29.	Lakshadweep	10	-	-	-	0.0%
30.	Mizoram	10	-	-	-	0.0%
31.	Pondicherry	10	02	-	02	20.0%
32.	Candidates from abroad	13	01	-	01	7.8%
Total		3,188	680	70	750	23.5%

REFERENCES

Arora, P.N. Number Appeared, Percentage Passed and Percentage of First Divisioners in Boards of Secondary and Higher Secondary Education in 1980 and 1981, NCERT, New Delhi, 1985.

Buch, M.B. Survey of Research in Education, Baroda: Society for Educational Research and Development (SERD), India, 1974.

Buch, M.B. Second Survey of Research in Education, Baroda: Society for Educational Research and Development (SERD), India, 1972-78.

Buch, M.B. Third Survey of Research in Education, Baroda: NCERT, New Delhi, India, 1985.

Bokil, S.R. Maharashtra State Board of Secondary Education, Pune, 1956-1963 in Second Survey of Research in Education, Baroda, SERD, 1974.

Comber, L.C. and John P. Keeves. Science Education in Nineteen Countries, Stockholm, Almquist and Wiksell, 1973.

Dave, R.H. "Factors Affecting Achievement", Third Indian Book on Education, NCERT, New Delhi, 1968.

Misra, R.G., Patel, P.M., Raina, M.K. and Mittal, J. P. Studies on National Talent Search, NCERT, New Delhi, 1981.

Passi, B.K. and Hooda, R.C. "Educational Evaluation and Examinations". In Third Survey of Research in Education, NCERT, New Delhi, 1985.

Passi, B.K. and Sansawal, D.N. "Educational Evaluation and Examinations", In Second Survey of Research in Education (SERD), Baroda, 1974.

NCERT, Review Committee on National Talent Search, NCERT, New Delhi, 1985.

Ministry of Education, The Secondary Education Commission, Ministry of Education, Govt. of India, New Delhi, 1952-53.

Ministry of Education, Education and National Development, Ministry of Education, Govt. of India, New Delhi, 1964-66.

Ministry of Education, Challenge of Education - A Policy Perspective, Ministry of Education, 1985.

Raina, M.K., and Mittal, J.P. Studies in National Talent Search, Section II, pp. 81-99, NCERT, New Delhi, 1981.

Shukla, Snehlata, " Achievements of Indian Children in Mother Tongue (Hindi) and Science", Comparative Education Review, June, 1974, 237-247.

The Assessment of Educational Achievement in the Indonesian School System

R. Ibrahim

I. INTRODUCTION

There are three major categories of examinations conducted in the Indonesian school system at present -- school examination, provincial examination and national examination.

1. School Examination

Internal School examination is undertaken for the purpose of improving teaching-learning process, reporting students progress, student classification and placement within each school, student promotion from one grade level to the other, and certification in certain subject areas.

In the primary and secondary schools, examinations include formative tests at the end of each unit of instruction, mid-term tests and end-of-term (summative) tests. Formative tests are used for improving teaching-learning processes whereas mid-term and summative tests are used for reporting students progress, classification and placement, and promotion from one grade level to the other. Besides, final examination for certification in subject areas like sports, arts, and practical skills, is also undertaken by each school.

In universities, examinations include mid-term and end-of-term tests for each subject matter, for obtaining course credits. For obtaining a degree at the end of the whole programme, examination is mostly undertaken in the form of oral exams with or without thesis.

2. Provincial Examination

Some provinces conduct regular examinations for Secondary Schools located in each province. Though the main purpose of these examinations is to assess and monitor curriculum implementation, the schools also use the results of these examinations in reporting students progress at the end of each semester.

Besides, each province also organizes examinations for participants of non-formal education programmes taking place in the province. These examinations are used for certification.

3. National Examination

There are three purposes for which national examinations are undertaken -- for certification, selection to higher level of education and monitoring educational progress. National examinations are used for primary and secondary schools only. In universities all examinations are internal examinations.

Further issues regarding national as well as school examinations/assessment will be discussed in the next section.

4. Notes

Besides being used for educational decisions, the results of assessment/examination are also used in selection for jobs in offices/companies. In developing its curriculum/examinations, the Directorate of Vocational Education always invite representatives from the Ministry of Manpower and various sectors in the society.

II. EXISTING SYSTEM OF ASSESSING EDUCATIONAL ACHIEVEMENT IN THE PRIMARY AND SECONDARY LEVEL

1. National Evaluation of Educational Achievement

There are several types of national evaluation undertaken to assess educational achievement in the primary and secondary schools -- National Final Examination, National Sampling Test and National Assessment of the Quality of Education.

a. National Final Examination

National final examination is an evaluation of student achievement at the end of primary, junior secondary and senior secondary schools throughout the country, undertaken since 1980.

This evaluation is undertaken to meet a number of purposes. The first purpose is for students' certification at the end of primary, junior secondary and senior secondary schools. In conjunction with this purpose, this evaluation serves the function of Leaving examination for which each province is given the opportunity to use local norms. The second purpose is for selecting students who want to continue their studies from primary to junior secondary and from junior secondary to senior secondary. In this respect, this evaluation also serves the function of entrance examination. The third purpose is for monitoring and improving the quality of education within the framework of current curricula. This includes the effort of reducing educational disparity among the provinces in the country. In this context, this evaluation serves the function of formative evaluation in its broadest sense. Since the use of national examination for monitoring may cause extra burden for the system, there is a need to plan a division of responsibility between regional and national authorities in this effort.

National final examination takes place once a year at the end of grade six of the primary school, grade three of the junior secondary school and grade three of the Senior Secondary School. This national examination obviously has significant effect on teaching-learning processes in schools. At present, this examination is undertaken by the Directorate General of Basic and Secondary Education (including the preparation of examination questions) in cooperation with the Testing Development Center. Since examination questions are constructed year by year, there is not much time available for validation of the tests. In the near future, the responsibility for preparing examination questions is going to be transferred the Testing Development Center. In assuming this new responsibility the Testing Development Center is now developing an item bank system for preparing better quality of examination questions.

b. National Sampling Test

National sampling test is an evaluation of student achievement within the primary, junior secondary and senior secondary, undertaken since 1979 to obtain information regarding the extent to which the objectives in the curriculum nave been achieved.

The main purpose of this evaluation is more or less similar to the third purpose of the National Final Examination, i.e. to monitor and improve student learning within the framework of the existing curricula. In other words, this evaluation is formative in nature.

This national sampling test takes place also once a year administered to schools which have been selected as samples from each province. In the primary school this evaluation is given to pupils at higher grade levels. In the junior as well as senior secondary schools, this evaluation is given to students at grades one, two and three. This evaluation has been undertaken by the Directorate General of Basic and Secondary Education since 1979 in cooperation with the Testing Development Center.

c. National Assessment of the Quality of Education

Since 1975, the office of Educational and Cultural Research and Development, in which the Testing Development Center is one of its Centers, has been undertaking the national assessment of educational quality.

There are three major purposes for which this assessment is undertaken. The first purpose is to obtain information regarding the national Mean of educational achievement as well as the disparity of achievement among the provinces/regions throughout the country. Using this information, the mapping of educational achievement throughout the country can be established. Recently this purpose has been extended to obtaining information regarding parts of the objectives which have been achieved and those which have not been achieved satisfactorily by the students. The second purpose is to examine the association between educational achievement and various input variables such as schools, teachers, facilities and the like. The third purpose is to find out how far some changes or progress in educational achievement through the improvement of inputs and process within a given period of time, have taken place.

This assessment is undertaken once every five years to schools which have been selected as samples from each province/region. There are three grade levels to which this assessment is undertaken -- grade six of the primary school, grade three of the junior secondary school (grade 9) and grade three of the senior secondary school (grade 12).

2. Ways of Assessing Student Achievement in Schools

As mentioned in Section I, internal examinations in the primary and secondary schools cover formative tests, mid-term tests and summative tests at the end of each term. Besides, final examinations for certain subject areas such as sports, arts, practical skills etc. are also undertaken by each school in spite of the fact that teachers' competence in undertaking assessment still becomes a problem.

a. Formative Tests

Formative tests are used for improving teaching-learning processes within each term. Accordingly, these tests are administered at the end of each unit of instruction.

Test items for formative tests are planned and written on the basis of instructional objectives to be achieved in each unit. Items should match instructional objectives in terms of both content and behaviour aspects.

In most cases, both objective and essay type questions are equally used by the teachers in preparing formative tests. With respect to the objective type questions, multiple-choice is the most popular form used by teachers. As to essay type questions, teachers tend to use restricted essay to make scoring procedure more objective and less complicated.

The results of formative tests are not presented in terms of total scores for each student. Instead, the results are presented in a way that the teacher could identify the items on which most of his/her students failed to answer correctly. This would enable the teacher to plan and provide remedial programmes for his/her students adequately. In spite of this advantages, the frequency and follow-up actions of formative tests have caused an extra burden for the teachers.

b. Mid-term and Summative Tests

Mid-term and Summative Test are used for grading and reporting students progress at the end of each term. Besides, these tests are also used for placement as well as for promotion from one grade level to the other.

Test items are planned and written on the basis of major instructional objectives covered in each term. Normally teachers establish a blue-print before preparing test questions.

Although essay type questions are also used in mid-term and summative tests, teachers tend to have more objective type questions, mostly multiple choice, in each test.

Test results are presented in terms of total scores of each student for grading and reporting purposes.

Schools in some areas also organize psychological testing covering the assessment of aptitude, interest and personality for use in student placement and guidance.

c. Final Examinations for Certain Subject Areas

Final examinations in certain subject areas serve the function of leaving examination for certification.

Teachers always establish a blue-print before preparing test questions which cover various topics and levels of cognitive ability to be tested.

In most cases, all the items represent objective type questions which can cover a long range of contents and can be easily scored.

As done in the mid-term and summative evaluation, test results are presented in terms of total scores of each student.

d. Note

In subject areas like Home Economics, Agriculture, Physical Education, etc. Moral Education Where psychomotor or affective domain is more emphasized, teachers also use practical tests or records to assess student achievement.

In using practical tests, teachers tend to evaluate both the product as well as the process through which the students accomplish the task assigned to them.

III. RESEARCHES/SURVEYS ON EDUCATIONAL ACHIEVEMENT

1. Major Researches Undertaken

There are a number of researches which have been and are being carried out on educational achievement in Indonesia, two of which may be regarded as major efforts i.e. National Assessment of the Quality of Indonesian Education and Study of Achievement in Written Composition.

a. National Assessment of the Quality of Indonesian Education

The National Assessment of the Quality of Indonesian Education conducted by the Office of Educational and Cultural Research and Development is a major evaluation study, consisting of three stages. The first is the Survey of Achievement in Grade Six which was first carried out at the end of 1975; the second, the Survey of Achievement in Grade Nine, conducted at the end of 1976; the third, the Survey of Achievement in Grade Twelve, was conducted toward the end of 1977.

This assessment study is Indonesia's first endeavor to examine carefully and critically the quality of learning at the various levels of schooling throughout the country. It is designed to provide information to the country's educational planners, decision makers, curriculum developers and researchers on what students have learned and on the various factors affecting the learning process.

b. Study of Achievement in Written Composition

This project is a part of the International Study of Achievement in Written Composition undertaken by the International Association for the Evaluation of Educational Achievement (IEA). Generally speaking the study is designed to find out ways in which written composition as part of the language instruction has been taught in schools and examine the level of achievement gained by the pupils in this written composition.

The international study covers 3 levels of education -- primary, junior secondary and senior secondary. As part of this international study, Indonesia has confined the scope of the national study starting in 1983/84 to primary school only due to some financial constraints.

2. Significant Findings

The study of Achievement in Written Composition is not completed yet so that the presentation of research findings would be confined to those obtained from the National Assessment of the Quality of Education.

It would be unreasonable to expect a large-scale cross-sectional survey to provide the polity-maker with a neat list of ready-made conclusions and recommendations for action. However, it is fair to expect a considered and cautious summary of the major findings which might have some bearing on educational policy.

Some of the major findings derived from the study reports are as follows:

a. Regional Differences

Extremely large regional differences exist in the levels of students' performance. This is equally true in each of the five major subjects of the curriculum: mathematics, science, social studies, Bahasa Indonesia and English.

Differences in the quality of teachers, text-books and other facilities, and financial situations seem to have significant contribution to differences in achievement.

This situation needs to be taken into consideration in making decisions in student certification based on national examinations.

b. School Libraries and Their Use

The presence of a library in a school, along with an adequate supply of books for borrowing by teachers or by students, is a variable strongly associated with achievement.

It stands out as clearly the most significant of a cluster of "School Quality" variables, and is independent of home background factors, in both urban and rural areas.

c. School Rolls and Class Sizes

Neither the size of the total school roll nor the size of the individual grade 6 and grade 9 classes is consistently related to student achievement, when other relevant factors are controlled.

In grade 12 study, however, size of school is a more powerful variable than it is at grade 9 as well as grade 6 studies, even after the effects of home background and school quality variables have been removed. This finding needs to be further examined and analyzed in more controlled settings.

d. Shared Facilities and Afternoon Schooling

There is an apparent advantage to students attending school in the morning, rather than in the afternoon (usually in shared facilities), but after the more powerful home background and school quality variables have exercised their influence, these two variables explain little further variation in total achievement. It appear that students attending in the afternoon are less able, since they were older and have therefore probably repeated a class.

Finally, it should be emphasized that the conclusions which have been drawn in this section can only point to potentially important areas for further investigation. Many of the conclusions will require further validation in experimental settings, and all of them will need wide discussion in order to spell out likely financial as well as educational implications. Besides, there have been also ideas for making use of the national final examination results for research purposes in the future.

References:

1. Report on the National Assessment of the Quality of Indonesian Education:
    a. Grade 9 (1978)
    b. Grade 12 (1981)
2. Laporan Penelitian National Kualitas Pendidikan tingkat Sekolah Dasar (1985)
3. Progress Report on the Study of Achievement in Written Composition (1985)
4. Concluding Country Report for the Fourth ASEAN Regional Conference on Test Development (1985)
5. Indonesian Country Report on Practical Test (1985)

The Assessment of Educational Achievement in Japan

Satoshi Kawanobe

Introduction

With the reform of the school education system after World War II, Japan adopted the 6-3-3-4 single-track system with 9 years of compulsory education. Six years in elementary school (Grades 1 to 6) and three years in lower secondary school (Grades 7 to 9) are compulsory, so that every child from 6 to 15 year old is entitled to receive education. In fact, 99.9% of children within the age group attend both elementary and lower secondary schools. In addition, more than 90% of 5 year old children now attend either kindergartens or nurseries before their entry into elementary schools.

There are three types of upper secondary school courses: full-time course for three years, a part-time course and a correspondent course, both for four years and more. In terms of teaching content upper secondary school courses are of two kinds: general, and specialized or vocational. Vocational courses include technical, business, agricultural, fishing, and others. Today, about 94.0% of the lower secondary school graduates (those who completed Grade 9 of compulsory education) go to upper secondary schools.

Education at university usually takes four years. Medicine and dentistry require six years of study. Junior colleges provide two or three years' education. About 38% of upper secondary school graduates now go to higher education such as university or junior college .

In Japan, there are no public external examination system during elementary and lower secondary education. Assessment of achievement for the purpose of deciding grade-to-grade promotion and graduation from school is made internally. The frequency of examinations for internal assessment varies from subject to subject as well as from teacher to teacher, but usually examinations for summative purposes are given twice during each school term (The Japanese school year consists of three terms). These two examinations are called 'mid-term examination' and 'term-end examination' respectively. The duration of examination periods is set by each school. In addition to these periodical examinations, tests for assessing the level of achievement and diagnosing teaching-learning deficiencies are given by teachers on an ad hoc basis during regular classroom sessions. Grade-to-grade promotion and graduation from the school are decided after taking into account the results of these examinations and tests together with the record of teachers' observations and pupil attendance.

Admission to upper secondary schools, colleges and universities, both public and private, are based on very competitive screening examinations. For public upper secondary schools the tests are administered by the prefectural or municipal boards of education which run the schools. Reports from lower secondary schools are also considered by each school when admitting students. In the case of private schools, entrance examinations are conducted by the individual schools.

The national and local public universities have two stages of screening for their 4 year undergraduate programmes. In the first stage, the applicants have to take the 'Nation-wide First-stage Entrance Examination' common to all national and local public universities. The objective is not to determine failure or pass, but to evaluate the overall basic achievement of the applicants. Results are sent to each university where the person has applied. In the second stage, entrance examinations are set by each university. Final admission is determined by putting together the outcomes of both examinations and using reports from each applicant's upper secondary school. In the case of the private universities, there is only one stage: the entrance examination given by each private university.

Although no external examination system exists for elementary and lower secondary education in Japan, various achievement tests are frequently conducted by national and regional authorities with a view to collecting data for the revision of the Courses of Study, and improving teaching materials and methods.

1. Research Studies and Surveys of Achievement

After World War II, more attention was paid to the achievement of children, and many educational administration agencies as well as educational research institutions conducted various surveys. These studies can be grouped as follows:

(1) Nation-wide achievement surveys conducted by the Ministry of Education, Science and Culture.

(2) International studies on achievement of mathematics and science initiated and coordinated by the International Association for the Evaluation of Educational Achievement, known as IEA.

(3) Surveys conducted by the National Institute for Educational Research (NIER) including its joint studies with the National Federation of Educational Research Institutes (NFERI) of Japan.

In addition, prefectural and municipal boards of education/education centres and voluntary educational research organizations with their own specific objectives from time to time conduct studies on the scholastic achievement of children under their jurisdiction.

(1) Nation-wide Achievement Surveys by the Ministry of Education

The Ministry of Education, Science and Culture conducted national surveys on the scholastic achievement of children in elementary, lower and upper secondary schools for eleven years from 1956 to 1966. The set objectives were to: a) formulate policies for the school curriculum standards and the improvement of instructional methods; b) take necessary measures for the improvement of learning conditions; and c) help individual schools improve their instructional activities by means of providing comparative data on achievement throughout the country. The surveys were conducted on the same day throughout the country using the same test items. Sample schools and grades within schools were randomly selected, and tests were administered to all the children in the designated classes. Table 1 shows the sample size, the grade level and subjects for each year.

Table 1: National Surveys on the Scholastic Achievement conducted by the Ministry of Education (1)

Year	Elementary			Lower Secondary			Upper Secondary		
	Subject	Grade	Sample	Subject	Grade	Sample	Subject	Grade	Sample
1956	Language Arithmetic	6	5%	Language Mathematics	3	5%	Language Mathematics	3	10%
1957	Social Studies Sciences	6	5%	Social Studies Science	3	5%	Social Studies Science	3	10%
1958	Music Arts Home-Making	6	5%	English Pre Vocational	3	5%	English Physical Education	3	10%
1959	Language Arithmetic	6	5%	Language Mathematics	3	5%	Language Mathematics	3	10%
1960	Social Studies Science	6	5%	Social Studies Science	3	5%	Social Studies Chemistry	3	10%
1961	Language Arithmetic	6	5%	Language Mathematics Social Studies Science English	2-3	100%	English	3	10%

Table 1: National Surveys on the Scholastic Achievement conducted by the Ministry of Education (2)

Year	Elementary			Lower Secondary			Upper Secondary		
	Subject	Grade	Sample	Subject	Grade	Sample	Subject	Grade	Sample
1962	Language Arithmetic	5-6	20%	Language Mathematics Social Studies Science English	2-3	100%	-	-	-
1963	Social Studies Science	5-6	20%	Language Mathematics Social Studies Science English	2-3	100%	-	-	-
1964	Language Arithmetic	5-6	20%	Language Mathematics Social Studies Science English	2-3	100%	-	-	-
1965	Social Studies Science	5-6	20%	Language Mathematics Social Studies Science English	2-3	100%	-	-	-
1966	Language Arithmetic Music	5	20%	Language Mathematics	1	20%	-	-	-
				Language Mathematics Pre Vocational	3	20%			

From 1961 the Ministry decided to survey all the second and third years' students of lower secondary schools in the designated subject areas. However, several criticisms about surveying the entire population were made including: the heavy burden on classroom teachers in conducting surveys; and the undesirable use of survey results to evaluate teachers' competencies. The Ministry decided to discontinue the survey project in 1967.

From 1982 to 1984, the Ministry of Education again conducted nation-wide surveys on the achievement of elementary and lower secondary school students to obtain basic data for improving curriculum standards. Subject areas surveyed and the size of samples are given in Table 2.

Table 2: Nation-wide Surveys on the Achievement conducted by the Ministry of Education from 1982 to 1984

Year	School	Subject	Grade	Sample size
1982	Elementary	Language; Arithmetic	5 & 6	16,000 each
1983	Elementary	Social Studies; Science	5 & 6	16,000 each
	Lower Secondary	Language; Mathematics	1 to 3	16,000 each
1984	Lower Secondary	Social Studies; Science; English	1 to 3	16,000 each

(2) International Studies on Mathematics and Science Achievement initiated by IEA

NIER acts as the Japanese national centre of the International Association for the Evaluation of Educational Achievement (IEA) and participated in its study projects on mathematics and science achievement. The objectives of the IEA studies are: (1) to study the state of instructional activities in certain subject areas such as science, mathematics, etc. in participating countries; (2) to examine achievement levels in each of the subject areas surveyed; and (3) to discover what factors affected achievement level. They make it possible for participating countries to analyze the results of studies in an international context, and to use them for the improvement of their respective situations. NIER so far has taken part in the first and second studies on mathematics and science achievement respectively. The relevant figures for Japan are given in Table 3.

Table 3: The First and Second Studies on Mathematics and Science Achievement Initiated by IEA

	Elementary			Lower Secondary			Upper Secondary		
	No. of Schools	No. of Pupils (Gr. 5)	No. of Teachers	No. of Schools	No. of Students (Gr.2 - Math) (Gr.3 - Sci)	No. of Teachers	No. of Schools	No. of Students (Gr. 3)	No. of Teachers
Math 1964	-	-	-	210	2,000	220	440	5,200	440
1980-81	-	-	-	200	8,000	200	200	8,000	200
Sci 1970	250	2,500	1,500	200	2,000	750	-	-	-
1983	220	7,900	550	200	7,600	300	190	6,500	130

Schools participated in the studies were selected using a stratified random sampling method based on type, location and size of schools.

(3) Studies on Scholastic Achievement Conducted by NIER and the National Federation of Educational Research Institutes

NIER's first study, in 1948 was a nation-wide study of basic skills of Japanese language, reading and writing, among 16,000 school children and adults of 15 to 64 years old.

The objective of this study was to find out the level of language proficiency possessed by ordinary Japanese people by analyzing language skills in reading and writing required for daily life. The findings of this study revealed that the language proficiency of the ordinary Japanese people, except some of the aged, was quite high in terms of reading and writing.

NIER's second study was conducted between 1952 and 1955. The objectives were to find out the actual state of scholastic achievement of school children and to clarify problems faced in teaching-learning situations. Target populations were children in Grade 6 of elementary schools and Grade 2 of lower secondary schools. Four subject areas; Japanese language, arithmetic (mathematics), science and social studies were covered as shown in Table 4.

Table 4: Nation-wide Studies Conducted by NIER

Year	Level	No. of Schools	No. of children			
			Japanese	Mathematics	Social Studies	Science
1952	Elementary	130	4,500	4,600	4,600	4,600
	Lower Secondary	125	5,000	5,000	9,000	5,000
1953	Elementary	150	6,000	6,000	6,000	6,000
	Lower Secondary	150	6,000	6,000	6,000	6,000
1954	Elementary	160	6,700	6,700	6,700	6,700
	Lower Secondary	160	6,700	6,700	6,700	6,700

After 1955 the conduct of regular nation-wide surveys on scholastic achievement was assumed to be a part of the Ministry of Education's responsibility. NIER still carried out nation-wide studies for limited purposes. One example is NIER's study of academic achievement attained by elementary and lower secondary school children, which was conducted from 1973 to 76. This study attempted to identify the kind and groups of items which were easy or difficult to answer for children from Grade 1 of elementary schools to Grade 3 of lower secondary schools including some items not yet taught at school. These were three subject areas; Japanese language, arithmetic (mathematics) and English. The sample schools were from Eastern Japan only and they were selected from urban, flat farming, and mountain farming areas.

From each of these three areas, four elementary schools were randomly selected, and the public lower secondary schools which admitted their graduates were also used in the study.

During 1975 to 78, as a part of the research project entitled 'A Comprehensive Basic Study on the Improvement of School Curriculum in the Perspective of Lifelong Education', NIER surveyed academic achievement and problem-solving ability among school children. The aim was to try to identify the capacity for 'self-educability' which is considered a prerequisite for survival in the rapidly changing contemporary society.

In the first year of the project, a survey measured overall achievement of children in five subject areas; Japanese language, arithmetic (mathematics), science, social studies and English. The sample consisted of 17,000 children in Grades 6, 9 (3rd year of lower secondary schools) and 11 (2nd year of upper secondary schools).

In the second and third years of the project, two kinds of surveys were conducted: Survey A assessed basic knowledge and skills; Survey B analyzed the development of problem-solving ability. The structure of the sample is shown in Table 5.

Table 5: The Structure of the Sample in 1976 - 78 Surveys

Year	Grade	Subject	Number of Children	
			Survey A	Survey B
1976	E-6	Social Studies	2,000	500
		Arithmetic	2,000	500
	LS-2	Social Studies	2,200	500
		Mathematics	2,200	500
		English	2,200	500
1977	LS-3	Social Studies	2,200	500
		Mathematics	2,200	500
		English	2,200	500
	US-2	Social Studies	1,600	500
		Mathematics	1,600	500
		English	1,600	500
	Total		22,000	5,500

Note: E: Elementary School
LS: Lower Secondary School
US: Upper Secondary School

In addition, case studies were conducted in 1978 to identify the formulation processes of children's learning abilities and to analyze factors affecting such processes. The case studies came from detailed analysis of classroom instruction in elementary social studies, and social studies and English at lower secondary school in four classes each in four elementary and lower secondary schools.

Since 1980, NIER has been studying the entrance examinations to upper secondary schools and academic achievement in elementary and lower secondary schools under a project called 'A Comprehensive Study on Academic Achievement'. This is both an NIER project and at the same time a joint study with NFERI. The aim is to identify the issues of scholastic achievement through the analysis of test items used in the entrance examinations to upper secondary schools and of teachers' assessment of the children's achievement level in elementary and lower secondary schools.

(4) Studies on Scholastic Achievement conducted by Other Administrative Authorities and Research Institutions

Prefectural and Municipal Boards of Education and Education Centres, and various public and private institutions for educational research have also studied scholastic achievement, and produced one hundred and thirty research reports in the past two years, covering a wide variety of situations.

The followings are a few examples of such studies:

A Study on the State of Elementary and Lower Secondary Schools: Arithmetic/Mathematics : by Shiga Prefectural Centre for Comprehensive Educational Research.

A Study on the Understanding of Children's Achievement: by Morioka City Municipal Institute for Educational Research.

2. Implications from the Results of Achievement Surveys

The research studies and surveys so far explained were conducted with different objectives, to various populations in terms of size and school or grade level, involving different subjects. It is difficult to generalize from such variety, but some broad trends can be identified.

(1) Level of Achievement

Although there are dangers in simply comparing scores from achievement tests in various countries, the results of IEA's Mathematics and Science Studies indicate that Japan is in the group of relatively high achievement among the participating countries.

Also the results of the achievement survey conducted by the Ministry of Education, Science and Culture in 1983 - 84 for lower secondary school students indicated the relatively high mean scores as shown in Table 6.

Table 6: Results of 1983 - 84 Survey on Lower Secondary Students' Achievement conducted by the Ministry of Education

Grade	Japanese Language	Social Studies	Mathematics	Science	English
1	72.1	58.5	60.3	67.4	73.0
2	69.6	56.9	62.1	62.7	56.8
3	65.7	64.5	70.7	70.1	63.3

Reporting these results, the Ministry of Education commented that when the mean scores of the common items for both the 1964-66 survey and the 1983-84 survey are compared, those of 1983-84 survey are generally higher than those of the 1964-66 survey.

(2) Difference among the Regions and within the Region

The differences among and between regions of Japan are relatively small. This may be explained by the following:

a) Educational contents to be taught at school are prescribed in the Courses of Study published by the Ministry of Education as a national standard and all the schools throughout the country are required to follow them, and textbooks are prepared on the basis of the Courses of Study;

b) Most of the teachers are well qualified due to the qualification requirements nationally set for teachers and provision of in-service training programmes; and

c) The Japanese language is the only national language used by all the people in Japan and it is used as a medium of instruction throughout the country.

In view of these, it can be said that diversity of achievement among the students' achievement may not be so wide in comparison with the one in other countries, but there still exist problems to be tackled in respect of the achievement, and some of these problems are presented as follows:

a) The existence of very little difference in achievement among the regions and of a narrow range of highest and lowest scores was revealed by the analysis of the Japanese data of IEA's Mathematics Study. Even though the difference in achievement is very small among the regions, there is a tendency to have slightly higher scores in the urban areas. According to the achievement survey conducted by the Ministry of Education, mean scores of children in the areas which have the population of more than 100 thousand are higher than those of children living in the villages having less than 8 thousand people by 9 to 11 points.

b) Also, depending on the grades and subjects, there are a group of students of so called "low achievers" or "inadequate achievers" who could gain only the scores of less than 50%. In general, the proportion of such students are increased in higher grades. For example, the results of the Ministry of Education indicates that in elementary schools about 10% to 20% of pupils in each subject are classified as such category, whereas in lower secondary schools 40% in English and 20% to 40% in mathematics, social studies and science are fallen into such category.

c) Further, a joint study of the National Federation of Educational Research Institutes (NFERI) revealed that 33% of elementary school teachers think that more than three quarters of pupils could attain the objectives stated in the Course of Study, whereas only 17% and 15% of lower and upper secondary school teachers think like that. On the contrary, 22% of elementary, 35% of lower secondary and 49% of upper secondary school teachers think that one-third or only one-fourth of students could attain instructional objectives stated in the Courses of Study.

(3) Problems in Each Subject

The achievement level at the terminal year of compulsory education, i.e. 15 years old, can be measured through the analysis of the results of entrance examinations to upper secondary schools. From the results of such analysis the following can be pointed out:

Japanese Language :relatively high scores in basic items such as reading and writing of Chinese characters (Kanji), but low scores in comprehension of texts and phrases.

Mathematics :high achievement in computation, but low achievement in logical thinking, application and analytical abilities.

Science :high scores for multiple-choice type of items, but low for the items which require or experiments. Also, low scores for items requiring the application of scientific thinking abilities.

English :high scores in listening and pronunciation but low scores in grammar and English composition.

3. Use of Results for the Improvement of Administration and Teaching-Learning Practice

Generally, the results of studies and surveys are reflected in the national and local educational administration. Also they are made available to schools and teachers for the improvement of day-to-day teaching.

The Ministry of Education analyzes the results of achievement survey by subject, grade and problem areas or in relation to certain educational conditions. The material is used for curriculum revision which generally take place every ten years, and also for the formulation of necessary measures to improve such educational conditions as class size, provision of instructional materials and facilities.

The prefectural or municipal boards of education also make use of the results for the improvement of in-service training programmes, recruitment of new teachers, and provision of educational facilities for schools and students under their supervision.

For some time after the implementation of the national achievement survey in 1961, the results were used to make a rank order of prefectures based on the scores. Then, in some cases, special supplementary lessons were given before the survey in order to achieve better scores. The results were also used to assess teachers for relocation purposes. These uses led to negative attitudes among teachers towards the national achievement survey and teachers became reluctant to support them.

As a consequence, most of the studies including those of IEA, NIER and NFERI have narrowed their scope and adopted very specific limited objectives and target populations even if nation-wide.

The results in NIER's studies are published as NIER's research bulletins or reports and distributed to universities with education faculties and educational research institutes or centres throughout the country. Similarly, NFERI's research reports used as reference materials for in-service teacher training programmes so that classroom teachers can make use of the results for the improvement of their teaching.

Problems in Assessment

1. Throughout elementary and secondary education, in addition to 'mid-term' and 'term-end' examinations, tests are

frequently administered by classroom teachers to trace the learning process of children, to find out the exact achievement level of each child, and for diagnostic purposes.

Criticisms are often made about these teacher-administered examinations. First, they tend to put too much weight on assessment of cognitive domain, neglecting the psychomotor and affective domains. Second, they tend to become examination for the sake of the entrance preparation, and appeal most to students in high grades, especially prospective applicants for the next stage schools.

2. In case of public upper secondary schools, each local board of education prepares its own examination on the basis of the educational content outlined by the Course of Study. The present selection system decides pass or failure on the basis of just this single entrance examination.

   One criticism of this system points out that the pressures to prepare for these entrance examinations distort the normal learning procedure at lower secondary schools. In order to improve the present system, the Ministry of Education has suggested alternative procedures, for example: give students more choice; make recommendation-based admission much broader and more general; and give individual schools choices on screening procedures. Local boards of education are now examining these measures proposed by the Ministry of Education.

3. The over-heated college/university entrance examinations have become a social issue and provoked many public calls for reform. The National Council on Educational Reform has suggested several measures to improve current university entrance examination practice, some of which are, for example, (1) replacement of current nation-wide first-stage entrance examination by common test scheme to which private universities be also invited to use; (2) establishment/strengthening of admission office in individual universities; (3) improvement of career guidance; (4) provision of alternative chances for entrance examination to national universities; and (5) positive discrimination to graduates from vocational upper secondary schools.

4. More specific studies on achievement to meet specific needs of individual communities and schools may be required in addition to nation-wide and international comparative studies on achievement. Particular focus should be given learning motives, attitudes and process of individual children in relation to their achievement.

Assessment of Educational Achievement
- The Malaysian Experience -

Mohd Rosli Mohd Ghazali

Introduction

Although the term assessment has often been used interchangeably with the terms evaluation and measurement, when used precisely, it has a narrower meaning than evaluation but a broader meaning than measurement. Evaluation is concerned with decision-making and is value-based. Assessment is concerned with describing a performance, behavioral change or any activity of a pupil which can be observed or reacted to. This description of pupil attainment should be value-free as far as possible. The relationship between these terms is that, assessment precedes the final decision-making stage in evaluation and measurement is the process by which assessments are made. Evaluation on the other hand may be made with or without using the evidence from assessment.

The term educational achievement can have a broad meaning; much depends on what one infers as "educational". If "educational" refers to any behavioural change in a pupil, then all types of formal and informal learning have to be considered for the assessment. This would involve a massive assessment exercise especially when conducted on a national scale. On the other hand, if "educational" implies only the formal aspects in schooling, then assessing non-scholastic and informal learning such as those in the affective domain would not be taken into consideration.

In the Malaysian context the latter meaning seems to be the order and close to prevailing situation.

This paper describes the current practice of assessment in the country's school system. It begins with a description of the school system, the system of assessment, related issues and finally some aspects of research in assessing educational achievement.

The School System

The education system in Malaysia operates within the 6-3-2-2 framework. A child starts formal schooling at the age of six plus undergoing a six-year primary education, followed by a three-year lower secondary education, a two-year upper secondary education and a further two-year of post-secondary or pre-university education.

Primary education is free but not compulsory. Universal education is for nine years i.e. Grade through Grade 9, where automatic promotion prevails. Continued schooling into Grade 10 will depend upon success in the public examination at the end of Grade 9. Similarly entry into post-secondary education is essentially based on the attainment of a terminal assessment at the end of Grade 11.

In the primary school, three types of schools according to the medium of instruction are provided namely, Malay (National Language), Chinese or Tamil medium. Parents could enrol a child to attend any one of these schools. However all schools follow the same curriculum through the common-content syllabus. This is to ensure that the children of diverse ethnic groups have a common outlook and understanding, an essential element in fostering national unity.

On completion of their primary education at the age of eleven plus, pupils undergo a three-year lower secondary education. Here, the medium of instruction is Malay with English as the second language. For pupils from the Chinese or Tamil media schools, they have an additional year of study in the Remove Class before proceeding to Grade 7. This would enable them to improve proficiency in the National Language.

The lower secondary level provides a comprehensive type of education. Besides academic subjects taught at this level, elective subjects of pre-vocational nature such as Industrial Arts, Home Science, Agricultural Science and Commerce are offered. Promotion is automatic from grade to grade. However at the end of Grade 9, the pupils will sit for a public examination. Those who are successful will be selected to proceed to upper secondary education.

About 60+% of the Grade 9 pupils will continue into the two-year upper secondary education. At this level the pupils will be channelled into the academic stream of either Arts or Science, and Technical or Vocational streams. At the end of their second year another public examination will be administered, the results of which will determine continuation into post-secondary education (such as pre-university classes, diploma or college level courses) or into the world of work.

The pre-university education provides a two-year preparatory education for tertiary level. However entrance for a university education depends on the success in the public examination at the end of the second year.

The System of Assessment

Assessing educational achievement of pupils in Malaysian schools *falls into two categories, i.e. internal and external* assessments.

Internal assessment is school-based. This is done through periodic testing or examinations in school, normally conducted monthly, termly, half-yearly or annually. Some are formal assessments while others are not. These assessments are part and parcel of the school evaluative system and include both formative and summative types. The purpose of these formal assessments are (1) to measure pupils' performance in order to monitor their progress (2) to inform parents their children's progress. In addition, the feedback from these assessments will enable the teachers to plan strategies for the improvement of teaching-learning process in the classroom.

Apart from the formal assessments, schools also monitor and assess from time to time the pupil's physical development as well as their career interest through informal means for guidance and counselling purposes. Longitudinal profile on the pupils is maintained in a standardized performance known as Cumulative Cards.

In the New Primary School Curriculum (implemented in 1983) assessment procedures have been incorporated into teaching-learning process. Teachers are being stressed on the need for continuous assessment and covering non-cognitive aspects as well.

The types of instruments used to assess educational achievement in the schools tend to follow those of the public examinations. Hence, there are objective, essay as well as practical tests. However the quality of instrumentation leaves much to be desired. Most teachers do not have the necessary exposure in the technicalities of test construction particularly item writing. Furthermore, these teacher-made tests do not go through stringent psychometric procedures simply because expertise in this field is lacking in schools.

An effort to overcome the shortcomings in the quality of test instruments is being made through pooling of resources by a number of schools within a district. Subject committees are formed comprising subject teachers from a number of schools, to build test items for use in schools in the district. Through this co-operative effort, the tests are better prepared and generally have much improved.

Schools have their own system of reporting pupils' progress in report cards. Each pupils has his own report card in which raw marks or percentages and/or grades are used to indicate their level of performance in the tests. Pupils' attendance and teacher's general remarks on the pupils' overall performance are noted down. These report cards are given to parents after each examination, depending on the practice of individual schools.

The external assessments, on the other hand, are much better *qualitatively. These formal* assessments which are summative in nature, are given at certain points in the pupils' school career. These are centrally conducted by the testing division of the Ministry of Education, i.e. the Examinations Syndicate, and the Malaysian Examinations Council.

The following are the external assessments administered at school level.

(a) Standard Five Assessment (SFA)*

The Standard Five Assessment introduced in 1967, is designed to measure pupils' attainment at the end of their fifth year of primary school. This examination is conducted in the medium of instruction of the schools (Malay, Chinese or Tamil). Regardless of the medium, all pupils are assessed by using common test instruments.

The main purpose of this examination is to provide:

(1) feedback to schools with regard to their pupils' performance in various subjects at the end of the fifth year of schooling.

(2) data to teachers on the strengths and weaknesses of their pupils so that remedial measures can be taken in the last year of their primary schools before they move on to (Grade 7) classes in the secondary schools.

The results of this examination is also being used as one of the criterion to select pupils for admission into residential secondary schools and science secondary school.

The subjects tested are Malay (National Language), English, Elementary Science, Mathematics, History and Geography. Apart from these, the Chinese and Tamil pupils are required to sit for Chinese Language and Tamil Language respectively. All subjects are tested by means of multiple-choice objective type.

(b) Lower Certificate of Education (LCE)

The Lower Certificate of Education examination, introduced in 1957, is the first accredition and selection examination in the school system. The examination is conducted at the end of lower secondary education (Grade 9).

---

* From 1988 the Standard Five Assessment will be replaced by another assessment at the end of primary education. It will be known as Primary School Achievement Test.

The main purpose of the examinations is to provide:

(1) terminal certification of success at the end of lower secondary education.

(2) data for selection and placement into upper secondary schools in Arts, Science, Technical or Vocational Streams.

The number of subjects offerable in the examination is 25 but a pupil may offer a minimum of 6 and a maximum of 8 subjects for a full certificate. The subjects include languages, social studies, mathematics, science and a number of elective/pre-vocational studies subjects.

Most of the tests for assessing pupils at this level are multiple-choice objective type. However for language subjects, essay tests are also administered in order to assess the expressive skills of the pupils. As for the pre-vocational subjects apart from multiple-choice tests, the pupils will also have to prepare an assignment in a coursework (practical assessment).

(c) Malaysia Certificate of Education (MCE)

The main purpose of the examination is for selection and certification at the Upper Secondary level. It is a terminal assessment for Grade 11 students in Arts, Science and Technical schools.

A wide range of subjects are offered for this assessment. These subjects are divided into groups such as Languages, Literature and Religious Knowledge, Social Studies, Mathematics, Sciences, Art and Crafts, Technical and Commercial. The pupils, however, may offer between 6 to 9 subjects taken from a specified subject groups, in order to be eligible for full certification.

Almost all subjects are tested by means of traditional essay type questions. However a mixed of essay type in one paper for the subject and multiple-choice objective types for the other has become the trend for assessing pupil achievement. In the Languages another element that is tested is oral ability. It is felt that the use of these oral tests encourages the development of oral skills in schools. As such, in spite of questionable reliability these tests may have as good measures of language ability, they are utilized for their educational value. The Science papers, besides having questions that are essay type and multiple-choice objective, also have practical tests. In the technical subjects, pupils are also required to sit for practical tests. In some other subjects part of the assessment is done through a coursework assignment.

(d) Malaysia Certificate of Vocational Education (MCVE)

The main purpose of this examination is to provide accreditation for those who have successfully completed a two-year education in the Vocational Secondary Schools. This examination also serves as an entry qualification into the Polytechnics.

The subjects offerable falls into two categories: (1) general subjects which include Languages, Social studies, Mathematics and Science, and (2) vocational courses which include Trade Engineering, Commerce, Agriculture and Home Science. A pupil has to offer the general subjects and at least one of the vocational courses.

Since this examination has the task of certifying the pupils' theoretical knowledge as well as hands-on ability in performing the psychomotor tasks in a particular trade or vocation, performance tests are given due emphasis. These are two types of performance tests. The first is the one-shot measurement of a pupils' ability working in simulated condition to the job situation. Independent external assessors are used to assess the pupils' performance. The second type is a series of tests conducted at intervals known as Phased Testing. The test at each point takes the form of either a complete project or a laboratory assignment. Assessing pupils' performance through this procedure gives a fairer measure of the pupils' specific ability in a particular vocational trade.

(e) Malaysia Higher School Certificate (MHSC)

This is a terminal and certification assessment procedure conducted at the end of Grade 13 by the Malaysian Examinations Council. The results of this examination form the basis for selection into the universities.

An open certification system is practised whereby pupil who have passed in any subject will be awarded a certificate. This is a departure from the group certification procedure used at the Grades 9 and 11.

Issues

The development of a centralized external assessment system as is practised today is a consequence of the national educational system. It evolved out of the need for a common and standardized assessment procedure in line with the objectives and aspirations laid down by the National Education Policy.

However, it is also felt centralized examinations causes unnecessary anxieties amongst parents, teachers and pupils. In a study on the Curriculum Development Centre, Beeby found that the national examination system played too important a role that it hinders curriculum development.* Too much emphasis was given to the teaching of examination subjects while non-examination subjects however essential, such as Civics, were not given sufficient emphasis.

There are also times when the objectives and uses of assessments have not been clearly understood. The results of the SFA have been used for selection and placement whereas they are intended for monitoring standards, such deviations between test intent and test use render the original objectives of these testing programmes obsolete.

As has been indicated earlier at the school level too, there has been apprehension over the quality of tests given to the pupils. Criticisms on poorly built classroom tests are quite frequent. Generally a good teacher should be able to produce good items. However, this is not easy to come by. With heavy teacher load a teacher will have very little time to prepare good test items. But in most cases, it is felt that most teachers do not have good training in testing. Consequently one dreads to think of the errors in judgement based on poorly built classroom tests.

There is also a need to improve the accredition procedures in the secondary school examinations. As it is, the LCE, MCE, and MCVE are based on "group certification". Pupils required not only to pass in a number of subjects but these subjects must be from a specified number of groups of subjects. There are cases where pupils have done well in individual subjects but failed to meet the group requirements. Such pupils are only given statements. This puts them in a disadvantaged position because the society has a tendency of judging pupils by their general performance as reflected by the certificate grades. An alternative system of certification, reporting only a pupil's individual subject grades, as used in MHSC is being considered for use in the other examinations.

In the present examination system, there has been an overemphasis on cognitive achievement of the pupils. The non-cognitive aspect, especially the affective ones, have been somewhat neglected. However this imbalance may be corrected when the new curriculum, both at the primary and secondary schools are fully implemented.

In concluding, although the existing assessment system may have its shortcomings, but generally there is evidence that it has played a vital role in improving the quality of learning and in increasing the achievement of pupils in the national educational system. Improvements in the future would entail a total approach whereby a pupil would be assessed not only his cognitive skills but also the other dimensions of his personality, aptitude and interest.

---

* Beeby, C. E, 'Evaluation of C.D.C.' (Unpublished report, Curriculum Development Centre, 1976).

Research on Educational Achievement

Research on educational achievement in my opinion, can be classified into three main areas, i.e. focussing on:

(a) the measurement used in assessment;

(b) the performance levels or standards of achievement; and

(c) the relationship between performance and social variables.

The first classification focuses on instruments used in the assessment. The study will try to answer questions about quality of tests, development of test types, the conduct of test administration and the scoring and marking of pupils responses. It is essentially a psychometrician's domain and the main task here is to ensure that quality is maintained when measuring the pupils' achievement, so that subsequent interpretation would be fair, valid and reliable.

The second area of research deals with standards or levels of educational achievement in the school system. The results of an assessment programme would provide some indications of the standards of education of the schools and serve as bench marks for improving pupils' performance. A study in this area would look for performance or achievement indices that would gauge the standards of education at any point in time.

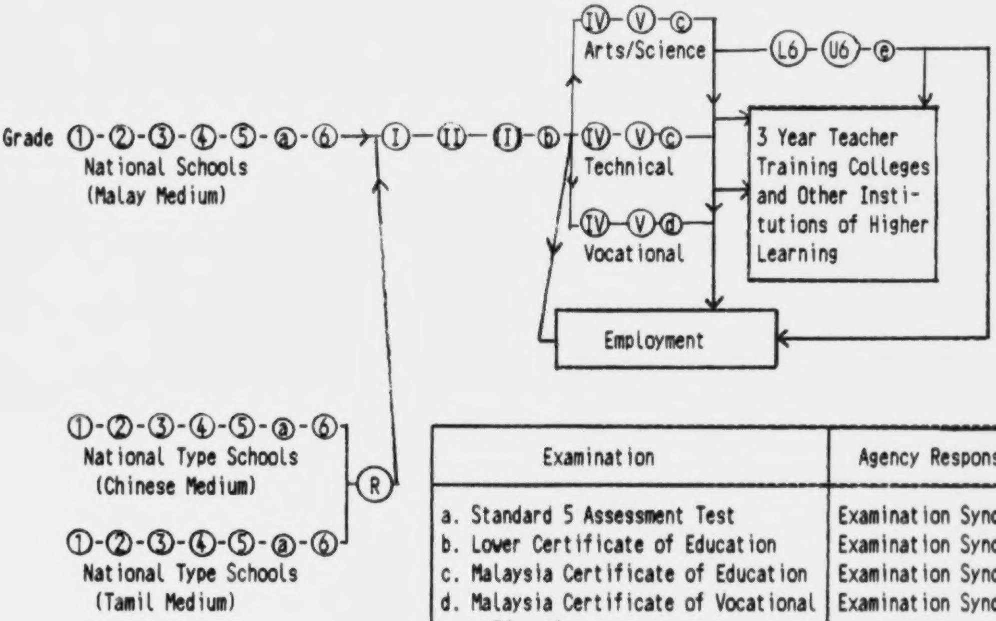
Focus of the third area of research is on the factors that may relate or contribute to pupils' achievement. This kind of study attempts to establish a causal relationship between pupils' achievement and any variables that may influence or relate to the pupils' performance. These variables may be student traits, teacher characteristics, socio-economic, school resources, school environment, etc.

Currently with the exception of a large scale monitoring exercise achievement of Grade 4 pupils which is to be conducted soon major research on educational achievement per se is virtually not in existent. However, there are research projects and monitoring programmes being conducted either institutionally or individually.

In the Malaysian context, research of the first type is mainly done at institutional level particularly at the testing division of the Ministry of Education. It is more of in-house on-going procedures to ensure that integrity and credibility of the institution is being maintained. At the same time the testing body gathers data for the second type of research activity. In most cases this is multi-disciplinary in approach whereby involvement from a number of institutions is practised. However there are cases where specialized institutions carry out their study for a very specific need for example, the constant monitoring of educational achievement in the New Primary School Curriculum by the Curriculum Development Centre. The third type of research has been mostly initiated by individuals rather than institutional. This is evident from the existing number of doctoral dissertations, done locally or abroad.

Appendix A: National Examinations in the School System

	Primary	Lower Secondary			Upper Secondary		Post Secondary	
Age	6 7 8 9 10 11 12	12	13	14	15	16	17	18
		13	14	15	16	17	18	19


Grade ①-②-③-④-⑤-ⓐ-⑥
National Schools
(Malay Medium)
Ⓘ—Ⓘ—Ⓘ-ⓑ
Ⓘ-Ⓥ-ⓒ
Arts/Science
Ⓛ6-Ⓤ6-ⓔ
Ⓘ-Ⓥ-ⓒ
Technical
Ⓘ-Ⓥ-ⓓ
Vocational
3 Year Teacher Training Colleges and Other Institutions of Higher Learning
Employment
①-②-③-④-⑤-ⓐ-⑥
National Type Schools
(Chinese Medium)
Ⓡ
①-②-③-④-⑤-ⓐ-⑥
National Type Schools
(Tamil Medium)

Examination	Agency Responsible
a. Standard 5 Assessment Test	Examination Syndicate
b. Lower Certificate of Education	Examination Syndicate
c. Malaysia Certificate of Education	Examination Syndicate
d. Malaysia Certificate of Vocational Education	Examination Syndicate
e. Malaysia Higher School Certificate	Malaysian Examination Council

123

Appendix B: Assessment in Malaysian Schools

Type of Assessment System	Grade Level	Purpose	Areas Assessed	Types of Instruments
Internal	All Grade Levels	Monitoring, Diagnosis, Selection, Placement	Mainly Cognitive, also include Psychomotor	Multiple-choice objective, essay, oral, non-formal assessment, performance
External	5	Semi Diagnostic, Achievement, Selection, Monitoring standards	Cognitive	Multiple-choice objective
	9	Certification, Selection, Placement, Monitoring standards	Cognitive, Psychomotor	Multiple-choice objective, essay, course-work
	11	Certification, Selection, Placement, Monitoring standards	Cognitive, Psychomotor	Multiple-choice objective, essay, oral practical, phase-testing
	13	Certification, Selection, Placement	Cognitive, Psychomotor	Multiple-choice objective, essay, oral practical

REFERENCES

1. Beeby, C. E. 'Evaluation of Curriculum Development Centre', Unpublished Report, Curriculum Development Centre, 1976
2. Kementerian Pelajaran Malaysia, Panduan Penilaian Dalam Bilik Darjah KBSR, 1983.
3. Mohd Rosli Mohd Ghazali, "Priority Needs of Test Development", Paper presented at National Seminar on Test Development 1981.
4. The Examinations Syndicate, "Report on the National Seminar on Test Development 1981", Ministry of Education.

Student Achievement, Evaluation System and Studies in Nepal

R.J. Thapa

1. The Educational System

The school level education in Nepal comprises of three levels. Primary (I-V grades), Lower Secondary (VI-VII grades), and Secondary (VIII-X grades). The objective of primary education is to impart literacy and to develop habits of good health and discipline among the students. The objective of lower secondary education is to develop character of the students and to promote respect for work among them. Secondary education aims at preparing citizens who could contribute to different areas of development of the country and to prepare students for higher education.

The Seventh Plan (1985-90) mentions that the structure of school level education will be changed over gradually to the 5+5 system, comprising of primary level (I-V grades) and secondary level (VI-X grades). At present, the final two grades of the ten year school education constitute the most crucial phase of school education, as the School Leaving Certificate (S.L.C.) Examination conducted by the Ministry of Education and Culture are based on the curriculum of these two grades.

Several national level agencies are involved in policy making, planning, and administration of education in the country. The National Planning Commission provides broad outline of national objectives and policy with respect to education. Another national level body, the National Education Committee, is engaged in policy planning and review. The Ministry of Education and Culture is the main agency engaged in the execution of education programmes. It has four major divisions: the Planning Division responsible for programming, budgeting, statistics and population education; the Educational Administration Division responsible for primary and secondary education, teachers training and adult education; Higher Education and Culture Division responsible for higher education, culture and scholarships, the General Administration Division, responsible for personnel and property administration. Besides, there are two important institutions, the Curriculum Textbooks and Supervision Development Centre, looks after, development of school level curriculum, preparation of school textbooks and technical aspects of Supervision, and the Controller of Examination Office looks after administration of the S.L.C. examinations. The Ministry also has two special project offices namely the Education for Rural Development (Seti project office and the Primary Education Project Office. There are five Regional Education Directorates in each of the five development regions of the country, and a district education office in each of the 75 districts of the country. The district education offices work in close coordination with district education committees, which have representation from the local institutions. Each school has its own managing committee, which is vested with wide powers relating to conducting of school affairs and teacher appointments and transfers.

2. The Development of Education in Recent Years

The development of education in the country has been remarkable during the past two decades. Progress in all aspects of education was specially notable during 1971-80 when comprehensive reforms were introduced in the education system under the National Education System Plan implemented since 1971. Among other things, this Plan introduced changes in the pupil evaluation system.

The rapid increase in population, hightended demand for education by the people, and liberal policy of the government in expanding provision of schooling facilities led to unprecedented quantitative growth of enrolments in all the three levels of school education. In 1984, the enrolments in different levels of school education were as follows: 1947 thousand at the primary level, 238 thousand at the lower secondary level, and 216 thousand at the secondary level. According to the estimates of the Seventh Plan, the gross enrolment ratios at the three levels in 1985 were: 78 percent at the primary level, 35 percent at the lower secondary level and 24 percent at the secondary level. The country is, thus, seen to be making big strides in attainment of universal provision of primary education. The representation of the lower secondary and secondary school age population in lower secondary and secondary schools has improved in recent years; still, a substantial proportions of this population is out of school.

Despite the several measures taken to improve the provision of education for larger number of students as well as the quality of education, the education system of the country is presently confronted with several problems. The problem of an observed decline in quality as reflected in high drop-out and repetition rates, and high rates of examination failures is frequently mentioned as a principal issue. The disparities in the educational facilities of the rural schools and the urban schools, and that of the public (government assisted) schools and the private schools have become glaring. The students of urban schools have shown better achievement levels than those of the rural schools. Similarly, students from the private schools have, in general, shown better achievement levels then those of the public schools.

Among other problems, the main ones are the low enrolment of the girls, shortage of qualified and trained teachers especially in the rural areas, lack of supplementary reading materials, increasing costs of providing education, and a general lacking of essential educational facilities and environment in most public schools. These problems are related to pupil evaluation systems in the schools. Pupil evaluation has tended to be made on traditional pattern of annual examination based mainly on essay type questions. And teacher teaching style is dominated largely by concern to prepare students for the set form of examinations.

3. Recent Educational Policy with Respect to School Level Education

The current educational policy of the government is reflected in the policy statements in the Seventh Plan. This Plan has put special emphasis on the fulfillment of the basic needs of the people. Primary education and skill based education have been identified as basic needs, among other things. The emphasis on the provision of free primary education as well as free primary textbooks has been reiterated in the Plan. The Plan aims at raising enrolments of all the three levels so that the gross involvement ration of 87 percent, 43 percent, and 33 percent at primary level, lower secondary level, and secondary level respectively could be reached by the end of the Plan period (that is 1990).

The strategies of education policy mentioned in the Plan include promotion of girls education, further development of technical education at the secondary school level (technical schools), upgrading of quality of education, fostering of competition among schools, strengthening of teaching of English, Mathematics and Science in the schools, development of physical facilities of the schools, and establishment of pre-primary schools. With respect to examinations, the educational policy mentions that the framing of questions, and the scoring of answer papers will be done in a scientific manner, and that activities like setting up of question bank will be taken up.

4. Quality of Education and the Examination System

There is a general concern that the quality of education has deteriorated in recent years. In Nepal, the quality of education is interpreted by different people in different ways. A school producing larger number of successful candidates in examinations (mainly the S.L.C.) is often said to have attained higher quality education. Common people also interpret fluency in English language achieved by a student as an indication of higher quality of education. Still others think that quality of education refers to ability of the students to speak and write Nepali and English language correctly and perform numerical calculations. Some people also align quality with the better discipline and behaviour of the students.

Quality of education is a broad concept, and covers several things: the academic performance of the student, the relevance of the content of education to the needs of the students and community, the kind of attitude and behaviour developed in the students by educational exposure. The examination system is used mainly to assess the cognitive attainment of the students. Other aspects of quality of education can not be assessed by conducting examinations, especially of the traditional type.

Lack of essential physical and educational facilities in the schools, low level of motivation of the students and teachers, lack of competent teachers in subjects of crucial importance like English, Mathematics and Science, and crowded classrooms are some of the factors responsible for poor quality instruction. Inferior instruction in the classrooms inevitably lead to poor examination performance of the students.

The examination system throughout all the levels has largely been a process of screening out students for higher levels of education. The system has not been used adequately for assessing the progress of the students' learning, and as a tool for improvement of the classroom instruction, the supervision system and the curriculum.

At present, the education system is heavily dominated by examinations. The teaching and learning activities are done primarily for the purpose of passing the examinations. The achievement of the schools and the performance of teachers are judged by the percentage of students passing the examinations. Examinations, thus, control the behavior, thoughts, attitudes of the students, teachers, and the educational authorities. The students have resorted to memorization, rather than trying to understand the underlying concepts.

Further, the students' (as well as parents') concern for success in examinations (especially the S.L.C. examination) has led to flourishing of private coaching. Students resort to private coaching both to remedy the deficiency in school teaching, and to ensure better performance in the tests. Many schools select and groom potential high performing students so that they (the schools) could score higher examination successes than other schools.

One of the urgent tasks of educational improvement is the upliftment of the standards of instruction in the schools. Provision of essential educational facilities, materials and trained teachers is essential for this. Continuous pupil evaluation system is also necessary. Examination at the one of the academic year is just a part of pupil evaluation. Evaluation could use various other means such as keeping of pupil's records, periodic tests, oral tests, assignments, participation in group discussions, etc.

5. The National Education System Plan Provision

The National Education System Plan introduced in 1971 tried to bring about fundamental changes in the evaluation system. The Plan's provisions with respect to evaluation were meant to remedy the defects of the examination system existing in pre-1971 days. The defects were identified as follows:

1) it was too much use of essay type questions (10 to 15 questions to be answered in 2 to 3 hours);

2) it led to cramming by the students;

3) it led to widespread copying and cheating by the students in the examination halls;

4) it was not able to make a true assessment of the students' learning; and

5) high failure rates were endemic.

The reforms introduced by the NESP were:

(a) Internal assessment system: The NESP stressed internal assessment as an important aspect of pupil evaluation. For this purpose, the test books of all the grades were made to include exercises and tests to assess the students' understanding of the lessons. Further, the NESP prescribed that 25 percent of the marks obtained in internal assessments at the secondary level would be added to the marks obtained in the final S.L.C. examination. This practice was terminated in 1980.

(b) Progress records: The NESP required the schools to maintain a progress record card for each student. Students would be given such progress reports among with the publication of school examination results. At present, most well run schools maintain the practice of giving such reports. However, many public schools have not been able to keep and provide such reports for lack of essential facilities.

(c) Objective tests: The NESP stressed the need of assessing the performance of the students in various subjects on an objective basis. Thus, it introduced objective type questions in the school level examinations as well as the S.L.C. examination. The NESP also emphasized the need of improving the method of scoring the answer books. The objective type question pattern prevailed for one full decade until it was abandoned in 1980. The difficulty of framing valid and reliable objective tests seem to have influenced largely the decision to abandon this system. At present short answers pattern (15 to 20 short answer questions to be completed in 3 hours) is in practice.

(d) External examinations: The NESP prescribed the following external examinations for different levels:

- District Level Examinations: Primary Level (end of the level examination).
- Zonal Level Examination: Lower Secondary Level (end of the level examination).
- National Level Examination: S.L.C. (end of the secondary level examination).

The district level examination of primary education was tried for some years. But it was discontinued particularly on the ground that it was not manageable to conduct a reasonably reliable test at this level. Thus, the primary level (end of the cycle, examination) is conducted by the schools concerned themselves.

The zonal level examination of the lower secondary level was also discontinued. Thus, the end of the cycle examination at the lower secondary level is also conducted by the schools concerned themselves. However, the government has again introduced this year a district level examination for Primary and Lower Secondary Schools. The examination is conducted by the district level Examination Committee headed by the District Education Officer. The examination is held throughout the Kingdom.

At the secondary level, the schools conduct qualifying examination, also called 'sent up test', for screening candidates for appearing in the S.L.C. examination. Most public schools have been lax in those sent up tests; and this partly explains the high rate of failures among students who have been 'sent up' to sit in the S.L.C. examination. The final S.L.C. examination (covering 7 subjects with total marks of 700) is conducted nation wide by the controller of examinations, under the Ministry of Education and Culture.

The subjects of study and marks alloted in examinations of the three levels: primary, lower secondary, and secondary (S.L.C. examination) is given in Appendix 1.

The NESP prescribed that a examination reform unit would be set up in the Controller of Examination Office to conduct studies on ways of improving the method of testing, and standard of question papers. It also recommended development of standardized tests. There is an examination reform unit in the Controller's Office.

The evaluation systems visualized by the NESP have became effective only partially. Well run schools have maintained student records and practised internal assessment to some extent.

Various factors such as lack of reference materials, non-availability of teaching aids, insufficient time to correct assignments, lack of essential facilities to provide remedial instruction to the students, etc. have hindered use of internal assessments.

Standardized tests have not been developed as yet. The process of constructing such tests has not gained momentum mainly due to lack of professional experts in this field, and limited understanding of the nature and implications of such tests in the country.

6. The S.L.C. Examination

The S.L.C. examination remains the ultimate test of proficiency gained by the students after being through the ten year cycle of school education. It is a test that makes or mars a student's future career, and job prospects (in the modern sector). Students passing the examination have direct access to higher education. S.L.C. qualification is rightly called a passport to higher education.

The S.L.C. examination is beset with several problems. One striking feature of this examination is the extremely high rates of failure or in converse, very low rate of successes. The average percentage of successful candidates in the S.L.C. examinations for 16 years from 1967 to 1982, is found to be 28.8 percent. The proportion of successful candidates reached its highest (53.2%) in 1969; since then it has gone down (with yearly fluctuation) to 16.7% in 1982. (Appendix 2)

Four different categories of candidates appear in the S.L.C. examinations in any particular year:

1) Regular (new) students,

2) The 'Exempted' (those who do not have to take school level 'sent up' test because they had failed the SLC test in previous years),

3) The compartmental students (those who are allowed to take the S.L.C. test in the subject or subjects they had failed in previous years),

4) The private candidates (those who have studied privately and have taken 'sent up' tests in schools).

The percentage of successful students among the regular students is much higher than in case of candidates of other categories.

The proportions of 'failed' students in S.L.C. examination of any particular year are staggering. It denotes inefficiency of the instruction in the schools. Most students are found to have failed in subjects like English, Mathematics, and Science. The weakness of teaching of these subjects is, thus, quite evident. The roots of the problem lie in shortage of competent teachers in these subjects in the schools.

There are several other problems of the S.L.C. examination system. The difficulty of management of the examination, involving ever greater number of the candidates is the principal problem. The Controller's Office is faced with the problem of shortages of manpower, and physical facilities to cope with the problem of managing the examination system.

Further, because the Controller's Office has been heavily involved in managing the logistics of the examination, it has not been possible to give enough attention to the important issue of improvement of the question papers and test administration system.

For improvement of the S.L.C. examination results, two things seem essential. Firstly, the quality of instruction in the schools should be improved. And secondly the system of examination itself should be improved.

7. Review of Research Studies on Achievement Levels and Examinations

Two studies of the achievement levels of the primary school completers have been conducted by the Centre for Educational Research, Innovation and Development (CERID) of Tribhuvan University in Nepal.

The first study carried out in 1980, entitled "Achievement Study of Primary School Children" came up with several conclusions with respect to performance of the three grade completers. The conclusions were based on the tests conducted by the study team. The study found that, of the total students taking the literacy tests, 20.4 percent were 'fully literate', 51.3 percent 'literate', 27 percent 'partially literate or 'semi-literate' and 1.2 percent 'illiterate'. Further, the proficiency gained by students in reading and writing was found to be satisfactory, but in arithmetic the proficiency was low. The study suggested that emphasis should be placed on drills to improve proficiency levels of the students in writing and arithmetic. The study also found the proficiency levels of the students in literacy test in the Far-Western Region of the country to be lower than that of the students in the Eastern Region.

The most recent study on the achievement levels of the primary school students is the "Effectiveness of Primary Education" study taken by the CERID in 1984 for the National Planning Commission. The findings of the study with respect to achievements of five grade completers are as follows: (by the subject of study)

- In Nepali, which is a major subject at the primary level, 57 percent of the students scored less than 50 percent marks in the tests.

- In Mathematics, 85 percent of the students scored less than 50 percent marks in the test.

- In Social Studies, 67 percent of the students scored less than 50 percent marks in the test.

- In English, 58 percent of the students scored less than 50 percent marks in the test.

These results of the tests on various subjects indicate that large number of students have not attained satisfactory level of achievements in various subjects. Achievement of most students is particularly low in Mathematics. The study found that most schools in the districts surveyed, faced shortage of teachers in Science, Mathematics and English. Teachers in these districts, indicated shortage of educational materials, inadequate physical facilities and non-attendance of the students as the main problems of the schools. Further details of this study is presented in Appendix 3.

No study of achievement levels of the secondary level students has been conducted as yet. However, the CERID had conducted a study of relationship of school facilities and the S.L.C. Examination results in 1981/82. This study compared to examination results attained by high performing schools (with 70% or above pass results in the S.L.C.) and by low performing schools (with 30% or less pass results in the SLC) and then related these results with the physical facilities of these respective schools.

This study finds that the high performing schools, on the whole, present a picture of a large size self-sustaining school with strict admission and promotion policies. These schools, in general, have adequate physical space as well as necessary instructional facilities and are staffed by an adequate number of teachers. Naturally all these schools are in sound financial conditions, and they have potential for further growth and development.

On the other hand, the low performing schools are faced with inadequate facilities, and funds which remain as stumbling block in adopting an effective instructional policy. The absence of strict admission and promotion policies, large size classroom (more than 60 students in a class), and the lack of instructional materials make classroom teaching unmanageable and ineffective in these schools. This, in turn, retards the development of enthusiasm and professionalism in the teachers. all these factors have adversely affected students' learning and achievement status. In brief, the low performing schools appeared to be caught in a vicious circle of inadequate facilities, inefficiency of instruction and poor quality of results in examinations.

In 1982 CERID also undertook a multi-pronged project entitled 'Instructional Improvement in Primary Schools' with a view to enhance the active participation and involvement of the teachers, the learners and the parents with a heavy emphasis on the development of essentials instructional materials for use by them.

The three-pronged approach is broadly classified into three-categories:

(a) Teacher preparation: The modalities adopted in teacher preparation included teachers workshop, subject teachers' meeting and individual consultations with teachers. In addition, teaching directions were prepared to facilitate efficient and meaningful teaching/learning activities in primary school classroom.

(b) Parental support: Parents of the children studying in the project schools were invited to a workshop session which focussed on the maps and means of facilitating children's education at home.

(c) Students' Learning Materials: Student work-books are supplementary nearby materials were developed in close consultation with teachers and other subject experts.

The analysis of the first year and second year achievement data revealed that the gain in students' achievement was significant in favour of project schools.

8. Some Measures for the Improvement of Testing System and Achievement Levels of Students in Nepal

Testing System:

(a) Internal Assessment: All schools should be made to adopt internal assessment system. Internal assessment conducted by the schools should be comprehensive and evaluate all aspects of student's growth and progress. Written examinations conducted by the schools of all levels should be improved. Teachers should be oriented in adopting new testing methods. Regular assignment or home-work should be given to the students. All schools should be made to maintain progress records.

(b) S.L.C. Examination: The S.L.C. examination should be improved by raising the technical competence of the paper setter, by improving the nature of questions by adopting scientific scoring procedures, and by using mechanical devices for the processing of the results. A pool of professional experts in tests should be developed.

(c) Improvement of Instruction: The instruction in the schools should be improved significantly, so that students could sit in examination with confidence of performing well. The S.L.C. examination results ultimately depend on the quality of instruction in the schools. Special attention should be given to improvement of instruction in English, Science and Mathematics.

(d) Involvement of Teachers: Teachers should be involved comprehensive pupil evaluation schemes. They should be trained in developing achievement, aptitude, and ability tests. Teacher training programmes should include contents on pupil evaluation system. The teachers should also be oriented to improvement of their questioning techniques in the classrooms.

(e) Development of Study Habits and Self Concept in the Students: The evaluation system should foster study habits and positive self concept in the students. Failure in examinations leads to negative perceptions of the self among the students. The method of evaluation should be such that it gives a reasonable feeling of possible success in the examinations. Regular assignments should be given to the students to encourage habits of study at home.

Achievement Level:

(a) Improvement of Instruction: School instructional facilities, provision of teachers, regularity of attendance of students and teachers in the schools, as well as physical facilities should be improved for this purpose. The public schools should put in special efforts in this direction, because they are particularly faced with poor quality of instruction.

(b) Improvement of Teachers' Morale and Professionalism: Teachers have crucial role in improving instruction. Informal assessment system is largely dependent on the teachers' integrity, as well as their proficiency in observing and assessing different aspects of the progress and growth of the students.

(c) Improvement of home backgrounds: The home backgrounds (level of living, educational level of parents, involvement of children in domestic work, overseeing of the children's study by the parents) substantially affect student's achievement. These conditions need to be improved by socio-economic development measures.

9. Recent Measures Taken by the Government

The government has recently taken following steps to enhance the efficiency of conducting the S.L.C. examination and for educational development.

- A new building for the controller office is being created in Sano Thimi educational complex near Kathmandu. This building with enough space for work, as well as maintenance of records will help promote the work efficiency of the office.

- Since past few years, the processing of the marks and the results is being done by computer. This has helped in reducing the interval between the time of the S.L.C. examination (held every year in winter), and the time of publication of the results.

Nepal is one of the member countries participating in the joint Innovation Project (of Unesco Regional Office, Bangkok) on 'Raising Achievement of Primary School Children'. The project is designed to assess differential effects of selected inputs (such as teacher preparation in materials development and evaluation techniques, parental involvement and supervisory help) on pupil in selected grades of primary schools.

The Joint Innovative Project (JIP) is being conducted in an integrated manner with the Primary Education Project (PEP) being implemented in six districts of the country with the assistance of the World Bank. The JIP activities will be implemented in the PEP districts through active participation of the District Education Office personnel, with the primary school supervisor (in the DEO office) playing a key role in the management and execution of the project activities.

The government is implementing two major innovative projects in the area of primary education: the Education for Rural Development (Seti) Project, and the Primary Education Project, mentioned above. Establishment of 'Resource Centre' and dissemination of improved instructional practices to the primary schools through these resource centres are the principal features of these projects. The aim is to enhance to effectiveness of classroom instruction and raise achievement levels of the students.

In view of the general ineffective instruction of Science and Mathematics subjects in the schools, the government has launched since 1983, a Science Education Development Project. Raising the quality of Science/Mathematics education, developing the competence of teachers, increasing the supply of Science/Mathematics teachers in lower secondary and secondary schools, improving the capability of the Ministry of Education and Culture (MOEC) in planning, management, monitoring, and evaluation of science education including curriculum development and increasing national capacities to design and manufacture science equipments are the objectives of this project. The project will undertaken in-service training of teachers, and supply equipment to 700 lower secondary and secondary schools, besides promoting science education at the college levels. A thrust given to improvement of science/mathematics education is certain to have very positive impact on the SLC results.

The Curriculum Textbooks, Supervision Development Centre of the MOEC is the main technical arm of the ministry. Its units on different subject areas are engaged in continuous study of trends in efficacy of instruction on different subjects at the classroom level, and also in devising ways of improving students' performance levels in the tests.

The government's recent policy of providing incentives (rewards) to better performing schools (in all the five development regions), and of fostering a spirit of competitiveness among the schools which allowing greater freedom in managing their own affairs and finances, is expected to make the schools put in more efforts to improve their instructional standards.

10. Concluding Remarks

Recently policy level personnel and educators in Nepal have become increasingly concerned with the qualitative aspects of the education system. In this context, the development and use of effective students evaluation has a very important role to play. Nepal has made considerable progress toward evolving different strategies of student evaluation. However, much remains to be done. The major aspects that need serious attention and immediate action are development of functional evaluation tools and proper reorientation of teachers in using the results of evaluation in instructional improvement. In this manner, evaluation can become an integral part of educational excellence.

Appendix 1

School Level Curriculum
(showing marks of the examination)

Primary Level Curriculum

Grade I - III

	Subjects	Full Marks
1.	Nepali Language	300
2.	Arithmetic	200
3.	Social Studies	100
4.	Physical Education, Hygiene, Handicraft, Drawing	50
	Total:	650

Grade IV - V

	Subjects	Full Marks
1.	Nepali Language	150
2.	Sanskrit	50
3.	English	100
4.	Mathematics	100
5.	Social Studies	100
6.	Science and Health	100
7.	Moral Education	50
8.	Physical Education	50
	Total:	700

Curriculum of Lower Secondary Level

Grade VI - VII

	Subjects	Full Marks
1.	Nepali	120
2.	Sanskrit	30
3.	English	100
4.	Mathematics	100
5.	Social Studies	100
6.	Science and Health	80 + 20
7.	Moral Education	50
8.	Physical Education	50
9.	Prevocational Subject	50
	Total:	700

Curriculum of Secondary Level

Grade VIII

	Subjects	Full Marks
1.	Nepali	100
2.	English	100
3.	Mathematics	100
4.	Vocational Subjects	100
5.	Science and Health	80 + 20
6.	History	75
7.	Geography	75
8.	Panchayat and Civic Life	50
	Total:	700

Curriculum of Grade IX and X

(Syllabus for the SLC examinations)

	Subjects	Full Marks
1.	Nepali	100
2.	English	100
3.	Mathematics	100
4.	Vocational Subjects*	100
5.	Optional A**	100
6.	Optional B***	100
7.	Extra Optional***	100

* One of the following subjects to be selected:

1. Home Science
2. Business Education
3. Office Management and Accounting
4. Agriculture
5. Trade and Industry
6. Cane and Bamboo Work (only for blind students)
7. Religious Rituals (for Sanskrit students)
8. Education.

** Two subjects from the following groups to be chosen:

1. Science Group: (a) Science,
   (b) Mathematics and Statistics.

2. Language Group:

   (a) Modern Language - Nepali, English, Russian, Spanish, Chinese, Japanese, Hindi, Bengali, Urdu. (At present, syllabus are prepared for Nepali, English, Hindi and Urdu only).

(b) Classical Language - Sanskrit, Greek, Latin, Arabic, Persian, Hebrew. (At present, syllabus is prepared for Sanskrit only).

(c) Local Language - Newari, Maithili, Tibetan. (At present, syllabus are prepared for Newari and Maithili only).

3. Social Studies Group:

(a) History, (b) Geography, (c) Civics, (d) Economics.

4. Commerce Group:

(a) Rural Economics, (b) Commercial Arithmetic, (c) Auditing, (d) Typing.

5. Sanskrit Group:

(a) Sanskrit Language, (b) Sanskrit Knowledge: only one (Justice, Grammar, Astrology, Literature, Vedas).

6. Home Science Group:

(a) Food Technology, (b) Clothing, (c) Child Care and Population Education.

7. Miscellaneous:

(a) Music, (b) Dance, (c) Architecture, (d) Painting, (e) Physical Education, (f) Health Education.

*** One subject from the selective groups mentioned above other than the ones selected as optional subjects to be selected

Source: MOEC, Curriculum, Textbook and Supervision Development Centre, Secondary Education Curriculum, Part I and II (in Nepali), 2039 (1982).

Appendix 2

Results of School Leaving Certificate (SLC) Examinations:

(Regular + Exempted + Compartmental + Private)

Year	Number Appeared	Number Passed	% Passed
1967	9,218	3,624	39.3
1968	12,134	5,871	48.4
1969	15,597	8,294	53.2
1970	17,192	5,271	30.7
1971	17,192	5,157	30.0
1972	15,772	5,267	33.4
1973	16,202	5,829	36.0
1974	21,286	7,813	36.7
1975	27,729	11,250	40.6
1976	32,720	7,208	22.0
1977	35,189	10,939	31.1
1978	41,008	13,961	34.0
1979	48,443	13,551	28.0
1980	63,989	18,789	29.4
1981	72,819	15,608	21.4
1982	79,968	13,337	16.7
	526,458	151,769	28.8

Source: MOEC, Educational Statistics Reports, Annual Series, 1977-1982.

Appendix 3

Some Details of the CERID Study on Achievement Level of Primary Level Students "Effectiveness of Primary Education"

1985

The study aimed at measuring the achievement levels of primary students at the end of grade V. Achievement tests were designed in four subjects namely Nepali, Mathematics, English and Social Studies on the basis of primary level curriculum and textbooks. The tests (of 50 marks in each subject) were administered to fifth grade students in sample schools of six districts Dang, Doti, Lalitpur, Panchthar, Mustang, Sarlahi Achievement levels in different subjects were as follows:

Nepali

The average achievement in this subject was found to be 23.32 with standard deviations of 9.44 points. The tests showed that the students had attained a relatively high achievement in comprehension and structure. However, their achievement in sentence construction, short answer items and full expressions was rather low. This indicates that classroom instruction had been oriented to getting the students to memorize and understand sentence structures and paragraphs in the textbooks, and not much attention was given to engage the students in sentence construction and language expressions. There findings strongly suggest the need of engaging the students in building up their vocabulary and in practicing composition and free expressions.

Mathematics

The achievement of primary schools students in Mathematics presented a very dismal picture. About 50 of the students taking the tests secured less than 50 percent of the total marks.

The students' achievement in solving word problems and four simple rules was relatively high. But the students had difficulty in solving problems that involve division and simplification. Only about 15 percent of the students could solve simple problems involving fractions and decimals. They were also found to have limited understanding of different units of measurement and their conversion. The students level of achievement in algebra expression and geometry was quite low. Students in some schools did not know how to use a protractor.

English

A majority of students (58%) got marks below the 50 percent level which shows that the overall achievement in English is below satisfactory level.

The analysis shows that the students secured a mean score of 24.05 with standard deviation of 12.96. The achievement of students in the vocabulary test in general was quite high (mean score 41). Though the students of Lalitpur and Mustang secured over 50 percent in structure, the overall performance in this section is rather poor (mean score 50%). The students performance in comprehension was about average (mean score 55%).

The remedial measures that could be taken to improve the situation lie in providing adequate training as well as reorientation training to primary school teachers.

Social Studies

In social studies, the students secured a mean of 22.26 (out of possible 50 marks) with a standard deviation of 7.43 points.

The analysis shows that the students performed well in social reasoning and skills. The majority of the students were found to have a fairly good understanding of life styles in different parts of the country. They also have good knowledge of heroes and builders of Nepal. However, they have some difficulty in answering questions dealing with migration, cottage industry and life-style of neighbouring countries. They also had difficulty in answering questions related to natural resources, maps and globe, religion, transport and the village panchayat.

The Assessment of Educational Achievement in New Zealand

Neil Reid

INTRODUCTION

Geographically, New Zealand is a remote group of islands tucked away in the south-west Pacific. And although its total area approximates that of Great Britain or Japan, just over three million people inhabit its two main islands, together with 70 million sheep. While most of the population is concentrated in three main urban centres, two in the North Island and one in the South Island, many people live in coastal sub-cities or rural towns.

New Zealand has had a short European history, compared with other nations of the Commonwealth; its Polynesian past is shrouded in myth and mystery. Its inhabitants are largely of British stock with a sizable minority (12% or so) of Polynesians - the native Maori and recent immigrants from many of the islands of the South Pacific. Increasingly, New Zealand is moving from a relatively homogeneous society toward a multi-cultural one, possessing its own individuality and an awareness of a separate destiny.

The strongly egalitarian temper of New Zealand society - a product of historical, geographical and economic factors - is reflected in education administration by the policy of making education freely available. Education as a right has long been prized and few other countries, it would appear, have achieved a closer approximation to equality of educational opportunity than has New Zealand.

Free and secular education is provided in State primary and secondary schools. Between the ages of 6 and 15 years education is compulsory for all children. However, the vast majority of children begin school on their fifth birthday and continue through to age 16, and many, particularly in urban centres, have had one or two years of pre-school experience before beginning their formal education. For those with ability and motivation, access is open to all branches of higher learning. University entry is currently possible at age 16 but increasingly more able students remain for an additional year (Form VII: modal age 17) of schooling before entering tertiary institutions or the workforce.

Assessment of Educational Achievement

The following diagram provides the major features of the New Zealand education system.

AGE

3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	2	1	2	STD. 1	STD. 2	STD. 3	STD. 4	FORM I	FORM II	FORM III	FORM IV	FORM V	FORM VI	FORM VII

←PRESCHOOL→ ←INFANT→ INTERMEDIATE S.C. U.E. U.E.S
KINDER-
GARDEN ←---------------- PRIMARY -------------------→ ←--------- SECONDARY ---------→

STANDARD/FORM: CLASS DESIGNATIONS
S.C.: SCHOOL CERTIFICATE EXAMINATION
U.E.: UNIVERSITY ENTRANCE EXAMINATION
U.E.S.: UNIVERSITY SCHOLARSHIP AND BURSARY EXAMINATIONS
▭: YEARS OF COMPULSORY SCHOOLING

The central administrative body directing education is the Department of Education, which is responsible to the Minister of Education. The Department's permanent head is the Director-General of Education; the head office is located in the nation's capital, Wellington. The Department also has regional offices in Auckland, Wellington and Christchurch, each under the control of a superintendent.

The control and management of State primary schools (about 2150 of them) is the responsibility of ten Education Boards and local Secondary School Boards control State secondary schools (some 300 in all).

There is also a substantial number of private schools conducted by religious bodies or private individuals. They are required to comply with certain standards under the Education Act 1964. Presently, many of these schools have integrated, or are contemplating integration with the State system.

ASSESSMENT OF EDUCATIONAL ACHIEVEMENT IN PRIMARY AND INTERMEDIATE SCHOOLS

The kinds and frequency of formal assessments to be made of children's educational achievement in New Zealand primary and intermediate schools are set out for principals in an official Department of Education publication, School Administration: A Guide for Primary and Intermediate School Principals (1975).

Section 9, under the broad heading of Evaluation, includes the following statement about the frequency of formal assessments:

> 9.7.1(b) On two occasions during the year, one of which shall be November or December, to review all available data about each child's achievements and record that assessment in the **Register of Progress and Achievement.**

These assessments are made on a 5-point scale (see p.5) and entered in the **Register**, the first assessment being made usually mid-year, in June, the second as specified.

Assessments are made in the areas of the curriculum listed below:

Language (oral and written)
Reading
Spelling
Writing
Arithmetic
Social Studies
Nature Study
Art and Craft
Music
Physical Education
Woodwork and Metalwork
Cookery
Dressmaking

In points to be noted by principals, it is stressed that evaluation is not to be equated with testing and that the criteria to be employed in making the assessments should reflect the aims and objectives of the school's programme and that these should be clearly defined and understood by the teachers using them.

Point (e) in Section 9.7.2, cautions principals that: 'Measurement or testing provides some of the evidence on which evaluation is based, but it should be clearly understood that interests and attitudes which form a significant aspect of the data on which the evaluation is made are seldom measured effectively by tests.'

However, in providing objective evidence for making formal assessments it is quite obvious that a good deal of information is to be derived from results of a wide variety of tests. Sections 9.7.3, 9.7.4, and 9.7.5 all deal with testing:

> 9.7.3 Evidence for Evaluation. The evidence on which an evaluation of a child's progress will be made will be gained, then, from tests of attainment, both standardized and teacher constructed; from the estimation of the pupil's potential as indicated by intelligence tests and observations of him in many situations, and an awareness of his level of development; from a knowledge of his attitudes and interests gained through daily observation; and through the careful recording of all this information so that it is readily available, not only for review and assessment, but also as a basis

for future programmes suited to children's needs.

9.7.4 Standardized Tests. These will include intelligence tests and tests of attainment. For assessing achievement, particularly in reading, standardized tests are widely used by teachers (for example, NZCER: Progressive Achievement Tests). It is the responsibility of the principal to see that these tests are administered correctly, interpreted wisely and used as the basis for the planning of subsequent programmes. The limitations of such tests (e.g., in the area on interest and attitudes) should always be fully recognized. Testing should always be purposeful; it is not an end itself; it should lead to better teaching.

9.7.5 Other Tests. Other testing carried out by the principal or class teacher will usually be predominantly diagnostic in nature. Its major aim is to indicate clearly to the teacher how effective his teaching has been, what each child's strengths and weaknesses are and what must be re-taught before proceeding to other topics in that particular field.

Several points made in these sections require elaboration.

(i) The use of the term 'intelligence test' has been discouraged in recent years; the notion of 'estimating potential' has been extensively modified and qualified in the literature on educational measurement.

(ii) The range of standardized tests available to New Zealand teachers has increased considerably since the publication of the guidelines for principals.

(iii) From observation, the 'other testing' conducted in classrooms by classroom teachers is rarely truly diagnostic but rather achievement testing of a formative or summative nature on specific topics. It ranges from informal, quick, 10-item quizzes to more extensive, formal, subject-orientated, hour long mini-examinations. Such testing may be done on an irregular or regular basis.

Current practice is extremely varied.

Mention was made earlier of the Register of Progress and Achievement and the formal assessment, or surveys made twice yearly and recorded using a national, norm-referenced, 5-point scale. Section 9.8 details their use by school principals.

9.8 REGISTER OF PROGRESS AND ACHIEVEMENT

9.8.1 General. The Register of Progress and Achievement is used to record the results of surveys conducted twice during the year, as required by the regulations.

The principal must have a thorough knowledge of the notes on the use of this **Register** (they are printed on the cover) and of the **Primary School Record Guide** card - E.19/23. It should be noted that:

(a) Surveys are required to be carried out twice yearly for all standard classes and for those infants likely to be promoted to Standard 1 at the end of current year.

(b) The survey is essentially a stock-taking exercise; it involves a review of all the data a teacher gains from day-to-day observation, and the results of formal and informal tests (including results in standardized tests of attainment and diagnostic tests).

Part II of the **Register** is specifically for classroom teacher use. Many teachers keep meticulous and useful records of student achievement in this section. Others make little or poor use of the space which has led some school inspectors to refer to it a 'white desert'!

9.8.2 Principal's Responsibility for Surveys. The principal has a particular responsibility for surveys. Final assessments on the 5-point scale as outlined in 9.8.3 below should be based on:

(a) A child's work during the half-year period.

(b) The results of regular tests carried out as part the class programme.

(c) The formal test, if any, by the head teacher or class teacher at the time of the survey.

(d) The teacher's knowledge of each pupil's aptitude and ability.

(e) Information obtained from standardized tests.

(f) Recorded notes of the day-to-day observation of the child, his interests and attitudes.

It should not be necessary to carry out a series of special tests at the time of the survey if an adequate programme of testing and evaluation in the various subjects is an integral part of the classroom programme. In accepting full responsibility for standards of work in the school, the principal must ensure that all formal tests are adequate for their purpose. He may, of course, require experienced teachers to prepare their own tests and submit them to him for approval, but he should see that inexperienced teachers are given help in the preparation of tests.

While the guidelines suggest that 'special tests' should not be necessary for survey purposes, many intermediate schools and large primary schools do conduct formal across-school testing at these times. The tests may be set by the principal, or senior teachers or a group of teachers at one class level who work together in a syndicate.

9.8.3 The Five-Point Scale. The five-point scale used at the time of the survey groups pupils into five broad categories: the work of an individual pupil is to be considered in relation to that of 100 representative children of his age (not of 100 pupils in the same school). The teacher is required to decide whether a pupil's attainments in an individual subject are, for his age group, average (3), above average (2), or below average (4). If an above-average pupil is judged to be so outstanding that he would be in the top five of one hundred representative children of his age he is given an assessment of 1. The below-average pupil who would be in the weakest five of one hundred children of his age is given an assessment of 5. The mid-year survey serves, among other things, as a trial and at the time of the second survey at the end of the year the first assessments are reviewed, amended or confirmed in the light of continuing observations and evaluations during the second half-year.

The back cover of the **Register** provides a table illustrating the assessment categories. Some teachers make further discriminations in the average group by using 3+ and 3-, thereby converting the scale to a 7-point one. A set of guidelines for teacher use also accompany the **Register** and spell out in some detail the criteria to be used in assessing educational achievement (among other things).

Quality of Performance	Outstanding	Above Average	Average	Below Average	Extremely Limited
Rating	1	2	3	4	5
Percentage receiving each rating	5	20	50	20	5

9.8.4 Reliability of Assessments. The principal is responsible for the reliability of assessments. The training of young teachers in the making of such judgements should be undertaken regularly. At the time of every survey there should be full staff discussion and guidance of individual teachers on the use of the criteria set out on the guide card (E.19/23) and in the **Register of Progress and Achievement** (i.e., the five-point scale and the educational objectives in each subject).

9.8.5 Survey Reports. After each survey the principal is required to write, in Part III of the **Register of Progress and Achievement**, a report based on the following criteria:

(a) An evaluation of the quality attained, standard reached, or progress made, in each area of the curriculum, in both basic and cultural subjects, for classes, groups, or individuals.

(b) Remedial measures needed to remedy weaknesses noted in surveys.

(c) Coverage of the curriculum and social attributes.

As the **Register** is a permanent school record, personal notes on pupils should be made carefully.

The final point should be noted. New Zealand now has an Official Information Act (1982) and school and teacher records can be accessed by both parents and the students themselves.

Because of some dissatisfaction with several aspects of the present form of evaluating and recording student progress and achievement, several schools throughout the country have been trialling a radically revised form of reporting for official purposes. In essence, the new system being employed experimentally abandons the norm-referenced basis of evaluation and encourages the provision of more qualitative and less quantitative data. The system also permits the purging of a student's personal file at various points in his/her school career.

ASSESSMENT OF EDUCATIONAL ACHIEVEMENT IN SECONDARY SCHOOLS

Secondary school principals have no official publication comparable to **School Administration: A Guide for Primary and Intermediate School Principals.** The secondary scene is characterized by its bewildering variety with no two schools assessing student achievement in exactly the same way or doing it with the same competency or comprehensiveness. The 'why', 'how', and 'when' of assessment are very much matters for individual teachers, subject departments, or syndicates of teachers to decide, although there are some hazy, broad patterns that can be discerned.

The writing of reports to parents on student achievement and progress usually entails a twice-yearly stock-taking exercise when information, both subjective and objective, is synthesized to provide a mark, grade, comment or broader statement on achievement within subjects. Additionally an overall written evaluation or assessment is usually provided by the Form teacher, dean or principal. Frequently, comments on student attitude, aptitude, interest, motivation, effort, etc., are also furnished for the information of parents.

The methods of reporting are just as variable as is what is recorded on them. Some schools are currently experimenting with structured student profiles in an effort to present a more balanced and rounded picture of student achievement and progress. But, many still simply record a mark, place in class and a teacher remark for each subject area.

Teacher mark-books and/or records of work completed will provide 'objective' evidence of individual achievement. These documents detail the results of formative and summative classroom tests (both oral and written), set assignments, practical work, aural tests, essays, projects, quizzes, formal examinations - in fact, the results of the whole gamut of assessment techniques which may, or may not, be appropriate. Herein lies a difficulty unrecognized by many teachers, resulting from the combining of scores and the condensation of information. Too often a distorted picture emerges when the conglomeration of scores with widely varying totals and standard deviations is added and averaged. Further, incorporated in the assessments being made of 'educational achievement' are part-marks (or consideration) given for neatness of presentation, effort, quality of illustrations/maps/diagrams, use of the English language, following of conventions or rules dictated by the teacher, etc., which all confound the measurement intended. Problems of weighting are also largely ignored by classroom teachers.

The frequency of assessment is also very variable. Some secondary school teachers test irregularly, some when a topic or unit of work has been completed, some on a regular weekly or monthly basis, others once a term in a comprehensive mini-examination, some in a cumulative fashion. There is considerable variation between subjects. Those that lend themselves to formal testing, such as mathematics and science, tend to be assessed more regularly. Subjects like social studies, economic studies and English, perhaps less so.

It is also fairly common for secondary schools not to set examinations in the third and fourth forms. Sixty minutes would probably be the longest time 13- and 14-year-olds would be required to persevere working on a test. It is not until students reach the senior school where external examinations are encountered that examinations of 2- and 3-hour duration are administered. Students would usually face two such exams in each of the fifth, sixth and seventh form years; one mid-year and the other prior to, and as a 'warm-up', practice exercise for one of the national external examinations. Some secondary schools, particularly the more academically oriented and independent schools, may set examinations at the end of each term (3 per year).

Unlike the typical primary and intermediate school where student records of achievement are readily accessible in the P **and A Register** and there are centralized systems of record keeping in place, secondary schools, being more complex institutions usually, leave much to be desired in this regard. Within subject departments record keeping may be very satisfactory and many teachers have comprehensive and meticulously kept files on their students' achievement. However, it is probably fair to state that a great many secondary schools presently do not have systems whereby a complete picture of a student's educational achievement could be gained from reading a single document at any given point in time.

What has been expressed above should not be interpreted as saying that thorough, technically and logically sound, and sensitive assessment does not occur in New Zealand secondary schools. That would be quite untrue: many schools make a concerted and sincere effort in this regard and some attain high levels of competency in providing valid, reliable and useful information on student achievement.

STANDARDIZED TESTS AND TESTING

Teachers have a variety of standardized tests in the basic skills, specifically developed for New Zealand, available to them. These, and other tests, notably of Australian, British or American origin, are purchased by schools and are used on a voluntary basis. The **Progressive Achievement Tests** series is administered extensively by primary and intermediate schools and to a lesser extent by secondary schools. Special bulk purchases of these tests are made from the publishers, NZCER, by the Department of Education through the Education Boards and regional offices. Following is an abbreviated description of New Zealand-devised or New Zealand-normed achievement and aptitude tests.

Progressive Achievement Tests

PAT: Reading Vocabulary: A test of words systematically chosen from the most commonly used 10,000 English words. The words are presented in the context of short sentences, with all questions in multiple-choice format. The child's task is to select the best synonym from five alternatives. Content-referenced scores and age percentile-ranks are provided for children aged 8-14 years inclusive. Administration time is 30 minutes.

PAT: Reading Comprehension: Short, graded prose passages representative of a New Zealand child's reading experience are presented to test both factual and inferential comprehension. Five to seven multiple-choice questions with five alternatives follow each passage. A unique feature of these tests is the presentation of level scores (content-referenced) related to the readability of the passages. These scores provide a valuable guide to the pupil's levels of competence which can be related to appropriate instructional material in reading. Suitable for children aged 8-14 inclusive. Administration time is 40 minutes.

PAT: Listening Comprehension: This test measures the child's ability to understand orally presented material of the kind commonly heard. Short, graded passages, both prose and poetry, are read by the teacher; the pupil's task is to select from among four alternatives the best answer to questions posed orally. Both receptive and reflective listening skills are measured and the results can be converted to age percentile ranks and level scores. In addition to helping teachers single out children whose listening skills are inadequately developed, these tests are useful in identifying those children whose listening comprehension is sufficiently different from their reading comprehension to suggest a remediable deficiency. The tests are suitable for children aged 7-14 years inclusive. Administration time is approximately 40 minutes.

PAT: Mathematics: A 50-item multiple-choice test designed to measure achievement in the important mathematical skills and understandings in each content area of the current New Zealand Department of Education syllabuses. At each class level, the questions are arranged roughly in order of difficulty in two cycles, so that a pupil attempts clusters of Knowledge, Computation, Understanding and Application items in each half of the test. Results may be converted to age and/or class percentile ranks. Suitable for children aged 8-14 years inclusive. Administration time is 40 minutes.

PAT: Study Skills: These tests measure skills in three broad curriculum areas not generally measured by other tests in the PAT series. These skill categories are: (i) knowledge and use of reference materials, (ii) reading and interpreting graphic and tabular materials, and (iii) general reading study skills (skimming, outlining, summarizing and the like). Three tests composed of varied stimulus materials typical of those encountered in both home and school contexts, are designed to tap each of the three broad categories of skills. Tests are available for children aged 9-15 inclusive for skill areas (i) and (ii) and 11-15 years for category (iii). Percentile-rank norms by age and class are provided. Administration times are approximately 40 minutes for each test.

Proof-Reading Tests of Spelling: The Proof-Reading Tests of Spelling (PRETOS) are broad measures of a child's ability to discriminate between misspelt words and correctly spelt words, presented in the context of meaningful paragraphs. The tests provide a measure of spelling achievement within the context of a proof-reading task, as well as giving diagnostic information about individual pupil's spelling accomplishments.

A separate non-overlapping test is provided for each class within the New Zealand range Standard 2 to Form II inclusive. Each test is made up of 3 or 4 paragraphs consisting of 12-14 lines of text, with 2 lines having no misspelt words. Class percentile ranks are provided for both production and recognition scores. The production score is a measure of the child's ability to spell a word correctly after identifying it as an error, and to locate the error-free lines. The recognition score is a measure of the ability to recognize misspelt words. Administration time is 30 minutes.

Burt Word Reading Test - New Zealand Revision: This widely used individually administered test provides a measure of an aspect of a child's word reading skills. Used in conjunction with other information, the Burt Word Reading Test provides useful indicators of possible wider reading problems. Normative data in the form of equivalent age bands are presented for 6- to 13-year-olds inclusive. Administration time is approximately 5 minutes.

Test of Scholastic Abilities: TOSCA is a group pencil-and-paper measure of a range of abilities involving manipulation of the verbal and numerical symbol systems familiar to English-speaking New Zealanders. The abilities tapped are school-related rather than curriculum-specific but generally do not involve skills of the kind taught directly in classrooms. Performance on the test will provide a guide to a student's current status in broad language and numerical reasoning abilities and give a broad indication of his standing in relation to age peers on the range of skills sampled.

Each test consists of a combination of 70 multiple-choice and completion items of widely ranging types, graded in difficulty. Three tests have been designed for different levels of the New Zealand school system (Primary, Intermediate and Secondary). New Zealand norms in the form of percentile ranks and stanines by age are provided for children from 9-14 years inclusive. Administration time is 30 minutes.

Standard Progressive Matrices - New Zealand Standardization: First published in 1938, the SPM has been used extensively in its original form. The test was standardized in 1984 for use in New Zealand. The SPM is a measure of non-verbal reasoning skills and should provide useful information complementary to verbal tests in any comprehensive assessment of children's abilities. Normative data by half-yearly age groups for New Zealand children between 8 to 15 years is provided in a New Zealand Norms Supplement. The test is untimed, but most children will complete it within 30 minutes.

The New Zealand Council for Educational Research has also recently adapted and normed the ACER: Advanced Test BL-BQ, ACER: Higher Test PL-PQ, and has collaborated with ACER in an Australasian adaptation and norming of the Differential Aptitude Tests. These tests and their supporting manuals, norms supplements, etc., will all be published and released to schools early in 1986.

In addition to the standardized tests described above, two other New Zealand-produced tests enjoy popularity: Marie Clay's **Sand: Concepts About Print** test for young children, and Clay, et. al's **Record of Oral Language**, also for primary-age children.

Several other measures are being developed or adapted by NZCER presently. Assessment aids for teachers in the field of written language (writing scales and interlinear, proof-reading tests of standard written English) are at various stages of development; the **ACER: Advanced Test B40** is being adapted; pre-entry diagnostic tests of mathematics were trialled in 1985 and the series may be extended and published in 1986; a cloze test of reading comprehension has been devised and trialled extensively and may be developed further this year. Staff of NZCER's Measurement and Evaluation Unit will also be experimenting with computer-adaptive testing in 1986 and beyond.

Other standardized achievement and scholastic aptitude tests, usually lacking New Zealand norms, that are used in limited numbers by New Zealand teachers are:

GAP Reading Comprehension Test
GAPADOL Reading Comprehension Test
Neale Analysis of Reading Ability
Standard Reading Tests
ACER: Silent Reading Tests
ACER: Arithmetic Tests
ACER: Intermediate Test D.

NATIONAL EXAMINATIONS

School Certificate

New Zealand school children do not face any national external examinations until they are well advanced into their secondary education.

The first of these examinations, School Certificate, is conducted by the Department of Education and is taken by the majority of students in November of their Form V year, that is, after 3 years of secondary education. The course of each candidate must include English, although the student is not required to sit the examination in that subject. A candidate may enter the examination in any number of the more than 30 options offered up to 6 and is credited with passes in individual subjects in which he or she is successful.

In 1984, 74,358 candidates entered 331,730 papers (this includes extramural and Pacific Island candidates). The most popular and widely taken examination subjects are English, mathematics, science and geography.

The examinations are prepared jointly by officers of the Department of Education and appointed examiners, who are practising teachers, under the general purview of the School Certificate Examination Board. The Board is broadly representative and includes members from the Post Primary Teachers Association, University Entrance Board, Department of Education, Independent Schools Association, Association of Heads of Independent Secondary Schools, and Ministerial nominees. The Board's functions are to advise the Minister and the Department of Education on matters of policy and to set in motion the annual preparation and execution of subject examinations for the awarding of School Certificate.

Papers are marked by panels of teachers hired for the purpose. They follow marking schedules prepared by the examiner(s). Checks are made on marker consistency/reliability using a variety of methods.

Scaling of marks then takes place. Within a subject, scaling is conducted to ensure reasonable comparability from year to year. A second more controversial scaling operation is also undertaken to establish a relativity or hierarchy among subjects based on a complex statistical procedure called 'means analysis'. The assumption underlying the employment of this adjustment is that students differing in 'general ability' choose different subjects. Hence, to be fair to the clever students taking 'difficult' subjects such as French and Physics, it has been determined that the pass rate should be higher than, say, for Human Biology or Home Economics which typically attract less able students. Pass rates in subjects therefore vary considerably. For 1984, for those subjects used as examples above: French 82.7%; Physics 82.6%; Human Biology 46.6%; Home Economics 44.9%;

Since its introduction in 1944, there have been a great many changes to the School Certificate examination. Some of these are worthy of note as they illustrate the refinements and modifications to what are essentially three-hour, summative, written measures of educational achievement in subject areas.

> In 1946, 65% of Form V students attempted School Certificate. Today 98% sit in at least one subject. Entries for the examination have grown from about 8,000 to almost 80,000 annually.
>
> Previously a mark of 200 or more had to be attained in 4 subjects, one of which must be English, in order to be awarded the Certificate. Now, single subject passes are awarded.
>
> Some subjects, namely Latin Studies, Art, modular Science, Geography, Indonesian, Japanese, Humanities, Agricultural Engineering and Workshop Technology are now wholly or par-

tially internally assessed in secondary schools. English is internally assessed in some schools that have opted to be included in the national 'trial' scheme. Several forms of moderation are employed for those schools involved with internal assessment.

Regional schemes have been introduced in some areas. These are locally organized, internally assessed and nationally moderated.

Marks are no longer recorded. Instead Pass or Fail and grades of A,B, and C are shown on the Certificate. A certificate is awarded to every candidate who gains 50% or more in at least one subject. No grades of D or E (Fail) are shown.

Increasingly, School Certificate as a national external examination has come under intense criticism, much of it ill-informed and unwarranted. Far too much is expected of the examination. A three-hour, content-oriented, paper-and-pencil test compiled by subject (but not test) specialists simply cannot meet such unrealistic expectations as those listed below:

Employers use S.C. as an aid to selecting staff.

It is required as a prerequisite for some jobs.

It is the basic for a number of Technical Institute courses.

S.C. is used as a control for entry to Form VI.

It is seen as a way of ensuring uniformity of national educational standards.

The community can use S.C. as a yardstick assessing the worth of teachers and schools.

Teachers use it as a means of motivating students to work hard from Form III onwards.

It is regarded as serving as a training for later, more difficult examinations.

S.C. is a means of ensuring fairness in the assessment of students.

The major criticisms (both valid and invalid) of the School Certificate examination are as follows:

S.C. enshrines the pass/fail concept instead of providing a statement of individual student achievement.

Acceptance of the pass/fail concept for S.C. guarantees that many students leave school with a sense of failure and no record of their accomplishments.

It is a 'one-shot' assessment and cannot adequately reflect a year's achievement.

S.C. measures a limited range of cognitive and affective abilities.

It restricts the development and flexibility of the school curriculum.

It imposes excessive emotional stress on adolescents.

S.C. dominates schools to the extent that they are compelled to organize for the needs of the examination candidates.

It is unfair to minority groups.

It discriminates against rural children.

It is over-emphasised in job selection.

Almost every shade of opinion exists concerning School certificate: from those who wish to return to past systems with 'tougher' standards, through those who are content to see continuing refinements to what is considered to be a proven system, to those who wish to see a major change in the nature of the examination. The flavour of the debate is encapsulated in the montage of recent newspaper articles appearing in one of Wellington's daily newspapers following the administration of the 1985 examinations. But, undoubtedly, if the present Labour government retains office in the next election, and the Minister's wishes are fulfilled, School certificate's days are numbered.

University Entrance

Students attending certain approved schools may be accredited for matriculation purposes provided they have completed a 4 years' secondary school course. Accrediting is conducted by Form VI teachers and is entirely internally assessed, the criterion being the student's ability to successfully undertake university studies.

Approved accrediting schools are required to justify their standards annually, by obtaining a specified number of passes in the U.E. examination. Students who are not accredited for entrance to university may quality by passing the University Entrance Examination which is conducted by the University Entrance Board. Since the bulk of those sitting the examination have failed to be accredited by their own teachers, pass rates are low, averaging around 25%.

The University Board is made up of representatives from the University Grants Committee, the universities, state secondary schools, independent schools, and the Department of Education. The Board has the statutory responsibility for prescribing syllabuses, appointing examiners, checking standards of accrediting and administering the examination.

Sixth-form students usually take 5 subjects, but may take 4 or rarely 6, from the more than 20 options available. Teachers consider performance overall in deciding whether or not to accredit a student.

The University Entrance Examination has also been subjected to fierce criticism. Major complaints have been:

> The requirements of the University Entrance have an undue influence over the School Certificate syllabus, even though most S.C. candidates will not take U.E. courses.
>
> University Entrance requirements have an undue influence on the Form VI curriculum, despite the fact that many sixth-formers do not intend going to university.
>
> The existence of two awards (U.E. and Sixth Form Certificate (see below)) is confusing and has had the effect of downgrading S.F.C. in the eyes of the public.
>
> Different bodies (S.C. Board, Department of Education, U.E. Board) controlling examinations in the senior school makes coordination and reform difficult.

The debate was cut short in 1984. A Labour government took office in July of that year. As part of its election campaign the Labour Party had pledged the removal of the University Entrance Examination, a move that had the support of the Post Primary Teachers Association (secondary teachers union), which had been clamouring for an end to the examination and a shift towards more school-based assessment.

The 40-year-old examination was conducted for the last time in 1985. Its abolition has opened the way for the Sixth Form Certificate to come into its own.

Sixth Form Certificate

With the increase in numbers in the sixth form during the 1960s, it became apparent that there was a need for a general leaving qualification which would have acceptability and standing in the community, but which would allow a more diverse range of courses and objectives to be assessed than was possible with the University Entrance qualification. The Six Form Certificate was introduced in 1969 to fill this gap and has co-existed with University Entrance in an uneasy relationship.

As originally introduced, the Certificate was awarded, not by the Universities Entrance Board, but by the Department of Education, to all students in New Zealand secondary schools who had satisfactorily completed a Form VI programme. The Certificate was internally assessed, without being moderated by a public examination or reference test. Each secondary school was instructed to assign grades to students on a five-point scale, according to a standard distribution in every subject: A - top 5 percent; B - next 20 percent; C - next 50 percent; D - next 20 percent; E - lowest 5 percent. A tolerance of 2 percent was allowed for each category, and larger deviations were sometimes permitted on request from schools.

After a trial period under this system it became clear that differences in levels of performance between schools were larger than had been anticipated, and that un-moderated grades resulted in a qualification which was not widely accepted in the community. One school's A was not equal to another school's A, and a desired parity of standards across schools was not being achieved.

The Sixth Form Certificate has undergone a series of modifications since that time. In its current form it is awarded by the Department of Education to all pupils in New Zealand schools satisfactorily completing a course in up to 6 approved subjects. It is necessary to take a course in English for 4 hours per week, but his subject does not have to be included in the subjects entered for Sixth Form Certificate.

In addition to subjects prescribed by the Universities Entrance Board for the University Entrance examination, some 200 'school-based' subjects are now offered by schools. The three most popular of these non-U.E. courses are typewriting, computer studies and physical or outdoor education.

Today, the Certificate is still based entirely on internal assessment within each subject, on a numerical grade system ranging from 1 (the highest) to 8 (currently the lowest grade awarded). The allocation of grades is determined by the performance of it non-repeating students in the previous year's School Certificate, which thus acts as a retrospective moderating device.

Each school receives from the Department a 'pool' of grades to be awarded to all students entering for Sixth Form Certificate in that school. The principal and Form VI teachers are responsible for allocating them to individual students in each of their subjects. The grade distribution is mandatory, but the school has complete freedom in dividing the grades amongst its students. Where a school considers it necessary to seek variation from its prescribed distribution of grades, an application may be made to the Department of Education outlining reasons for the variation sought, together with supporting evidence.

With the Minister of Education's promised axing of School Certificate by 1990, it will be interesting to see how Sixth Form Certificate will be moderated. A Committee of Enquiry, established by the Minister to look into the implications of the removal of the U.E. exam from Form VI, has recommended only that S.C. results be used in 1986 and that "the matter of alternative systems if inter-school moderation should be discussed more fully in the second stage of its deliberations on assessment in Forms V - VII" (p. 18).

Acceptance of the Sixth Form Certificate appears to be growing, aided by the removal of U.E. which overshadowed it for so long, and with the help of a vigorous publicity campaign mounted by the Department of Education and backed by the secondary teachers union (PPTA) and the Employers Federation.

The matter of entry to university, it appears, will remain the function of the Universities Entrance Board. In the absence of an examination the Board has decided that students with a Sixth Form Certificate award in one or more subjects may apply for provisional entrance to university. The university will be required to consider applications in terms of the student's fitness for university study in the degree or diploma in respect of which the application is made.

It is proposed that acceptance for entry to university from Form VI be based on a provisional qualification confirmed or otherwise by achievement at the university. This means that the Board must provide for the gaining of a definitive qualification from school at the end of Form VII. The Minister's committee, in addressing the broader issues set out in its terms of reference, will also make recommendations on assessment at Form VII and its findings will clearly influence the way the Board's award structure is eventually defined.

Higher School Certificate

Higher School Certificates are awarded after a 5-year course to students who have been accepted for entry to Form VI and have satisfactorily completed an advanced course for 2 years. H.S.C. is also awarded to students who have obtained a qualification recognized by the University Entrance Board for admission to a New Zealand university and satisfactorily completed an advanced course for 1 year.

University Bursaries and Scholarship Examinations

These are taken by more able Form VII students. They are competitive exams for supplementary awards for university study conducted by the University Entrance Board.

The examination and award picture is undoubtedly confusing and both curriculum and assessment in the senior secondary school are currently under Ministerial review.

The following table summarizes the range of awards now available:

Level and Award	Administered by	Main Features
[Form V]		
School Certificate	Department of Education through School Certificate Examination Board	External examination for most subjects, ranking, marks grades, pass/fail.
Local Certificates (Maths)	Professional associations of teachers	Various methods, mainly internally assessed.
[Form VI]		
Six Form Certificate	Department of Education	Ranking, grades, school-based, national moderation using S.C. results.
University Entrance*	Universities Entrance Board	External examination with accrediting; ranking, marks, pass/fail.
[Form VII]		
University Bursaries	Universities Entrance Board	External examination, ranking, marks, competitive.
Entrance Scholarships	Universities Entrance Board	External examination, ranking, marks, competitive.
Higher School Certificate	Department of Education	Course completion, non-graded.

* This examination not available after 1985.

MAJOR SURVEYS OF, OR RESEARCH INTO EDUCATIONAL ACHIEVEMENT

Apart from the regular checks on norms for New Zealand-standardized tests and the cross-national IEA studies (see below) there have been very few substantial surveys or research programmes focussing on educational achievement conducted in New Zealand.

Such surveys that have been undertaken have, almost without exception, been in response to public or political pressure. For example, in the early 70s when the heated debate over internal assessment was threatening to become incandescent, the timely publication of **External Examinations and Internal Assessments : Alternative Plans for Reform** (Elley and Livingstone, 1972) drew together the threads of what was known from overseas and New Zealand research on the issues and presented the findings in a scholarly and dispassionate manner. The book is still quoted from extensively, both in New Zealand and overseas and a method of moderating Sixth Form Certificate presented in the apppendices is still used by many secondary schools more than a decade later.

Two major surveys of educational achievement were conducted in 1978 in the light of growing disquiet over educational standards in New Zealand schools and in response to a 'back-to-basics' lobby which was becoming very vociferous at that time.

One survey conducted by NZCER (Reid, 1978) drew on all available New Zealand research information and collated and presented it in such a way as to illuminate the standards issue and, as an aspect of the debate, re-presented the case for national assessment in New Zealand.

The second survey, undertaken by officers of the Department of Education, was presented as a report to the Minister (Department of Education, 1978,). The report drew heavily on the opinions of inspectors of schools and other Departmental officers supported by objective evidence, where it was available.

The NZCER report found that:

> "...New Zealand educators do not have the reliable and objective information needed to answer the concerned layman's query: 'Are scholastic standards falling?' And the only way one can get that kind of data is by setting in motion extensive, well-planned and systematic national surveys to monitor what is happening in our schools" (p.9).

The Department's report in the final summary concluded:

> "The general impression that emerges from the review of standards is a mixed one" (p.98).

This was interpreted by many to mean that standards had generally been maintained over the past 25 years and that the educational achievements of most students were satisfactory although there were areas for improvement and some shortcomings in the system (see, "Reporting on Schools: Comments on 'Educational Standards in State Schools'", **Education**, Vol. 28, No.2, 1979).

A perusal of reports of current projects sponsored by the Department of Education reveals that, of the 46 currently in train, there are no major research exercises (apart from the two IEA studies) that have educational achievement as their major focus. There are several studies, however, such as Waikato University's 'Primary School Learning in Science Project', and the 'Social Studies Subject Survey' (Department of Education, 1984) that have evaluation of learning and/or educational achievement as an aspect of the research.

National Assessment of Educational Achievement

Despite recommendations from various committees, commissions and individuals, and proposals submitted to the Department of Education from institutions like NZCER, New Zealand has no system of national assessment comparable to, say, the British surveys conducted by the Assessment of Performance Unit or America's National Assessment of Educational Progress.

The closest equivalent is the regular monitoring of the performance of New Zealand children on the **Progressive Achievement Tests** conducted by NZCER's Measurement and Evaluation Unit.

Briefly, those schools selected as a representative sample that took part in the original standardization of the tests are invited to furnish data at 6- or 7- year intervals to enable the norms to be checked. On the first occasion, only the norms are checked. But both the norms and the item difficulty and discrimination indices are scrutinized on the second check. Results are reported in the first instance to the Council's broadly representative Advisory Committee on Educational Tests; later the findings are disseminated through the **NZCER: Newsletter** and the media.

A summary of the norms checks and findings is shown below:

PAT TEST	YEAR	GENERAL FINDINGS
Reading Vocabulary (Normed 1968)	1974	Slight decline in mean scores over all class levels.
Reading Comprehension (Normed 1968)	1974	Little change over all class levels; statistically significant increase in mean score at Std. 2.

PAT TEST	YEAR	GENERAL FINDINGS
Listening Comprehension (Normed 1970)	1977	Very small but consistent decline over all class levels.
Mathematics (Normed 1973)	1979	Small but consistent decline at all class levels; statistically significant at Form II level. (Secondary school data unusable)
Reading Vocabulary (Normed 1968)	1981	Tiny increase in mean achievement in Std. 2; declines of varying magnitude at other levels; Std. 4, Form I, Form IV all statistically significant.
Reading Comprehension (Normed 1968)	1981	Negligible gains at Std. 2 and Std. 3. Very small decline at other levels; Forms II and IV statistically significant.
Study Skills (Normed 1976)	1983	Virtually no change over all tests at all class levels.
Listening Comprehension (Normed 1970)	1984	Increasing decline in mean scores; statistically significant differences from Form I upwards. 3 raw-score points difference at From IV.
Mathematics (Normed 1973)	1985	Decline in mean scores over all class levels. 4 and 3 raw-score points at Forms III & IV.

The IEA Studies

New Zealand participated in the international surveys conducted by the International Association for Evaluation of Educational Achievement in 1970 and 1971. The target groups for these particular surveys in reading comprehension, literature, science and civics education were 14-year-olds (Form IV students) and 17-year-olds (Form VII students).

While cross-national studies of educational achievement lend themselves to rather invidious international comparisons, it is possible to obtain a very rough picture of student performance relative to others elsewhere, provided the numerous caveats, cautions and qualifications injected by the authors of the international reports are taken into account.

For reading comprehension, New Zealand stüdents at the age levels sampled compared very favourably with students in other countries, achieving the highest mean scores for Populations II and IV. However, on measures of word knowledge and speed of reading (for Population II only), they performed less well, achieving at about the median score for the 15 nations participating.

In knowledge of literature and responses to literature, New Zealand students out-performed students from other countries tested.

In a test of knowledge of civics, the New Zealand fourth formers achieved at about the international average. However, the Form VII student's knowledge was superior to that of their counterparts in 7 other countries tested for that part of the study. The authors of the international report characterized New Zealand students as being knowledgeable, participative and responsible individuals.

For science, the New Zealand sample for Population II achieved better than students from all other participating countries except Japan, Hungary, and Australia. The Population IV students, however, achieved better than those of the other 18 countries in the international study. New Zealand students performed creditably not only in recalling factual information, but also in the skills of applying scientific understanding in novel situations. This ability was much less evident among students in many other countries.

For the French as a foreign language study, conducted in only 8 countries, New Zealand children at the two levels tested (Populations II and IV) did rather less well.

More recently, New Zealand has participated in two further IEA studies, the Second Mathematics Study, conducted in 1981, and the Written Composition Study undertaken in 1984.

Complete data of a comparative nature for the former study, in which national samples of 13-year-olds (Form III) and 17-year-olds (Form VII) were involved, have not been released in widely accessible form as yet. The findings are not good, however, for Form III students and there have been repercussions from these results already. The performance of Form VII students was much more pleasing, apparently, with achievement at an acceptably high level.

Some very preliminary observations are available from the written composition exercises which were attempted by samples of 12-year-old (Form II) and 15-year-old (Form V) students, but they are not of a cross-national, comparative nature, and deal with student perceptions of the writing task, choice of writing topic, their likes and dislikes, and the types of error being made in spelling and grammar.

CONCLUSION

Quite plainly, major aspects of New Zealand's education system relating to educational achievement are in the throes, or on the threshold of sweeping changes. Ministerial reviews are in train in both the curriculum and assessment; a system of evaluating and reporting on student progress and achievement in primary and intermediate schools which has served for 20 years or more, is likely to be extensively revised; the 40-year-old University Entrance Examination has been banished from Form VI, and the School Certificate Examination in its present form is under threat of abolition by the end of the decade. There is also the beginnings of a ground-swell of opinion that will see teachers drift away from norm-referenced assessment to criterion- or objective-referenced educational measurement, accompanied by an acceptance of the concepts underpinning 'mastery learning', and of notions of 'schools without failure'. Further, in some quarters, there has been an elaboration of the prevailing doctrine of 'equality of educational opportunity' to 'equality of educational results' to be attained by a more equitable sharing of educational resources.

Whether these major modifications to the education system are wise and in the best interests of New Zealand society only time will tell. Certainly, they are not being made without a substantial measure of professional and public opposition.

References

Cole Catley, C., McAlpine, D.C. et. al. (1979). 'Reporting on Schools'. **Education**, 28:2, pp. 2-13.

Department of Education (1975). **School Administration: A Guide for Primary and Intermediate School Principals**. Wellington: Government Printer.

Department of Education (1978). Educational Standards in State Schools. Wellington: Government Printer.

Department of Education (1985). 'Looking at Learning in Schools, Issue 7, Assessment and Awards in the Senior Secondary School'. Discussion document, Ministerial Reviews: Curriculum and Assessment. Wellington Government Printer.

Department of Education (1985). 'Implications of the Removal of the University Entrance Examination from Form 6'. First Report of the Committee of Enquiry in Curriculum, Assessment, and Qualifications in Forms 5 to 7. Wellington: Government Printer.

Elley, W.B. and Livingstone, I.D. (1972). **External Examinations and Internal Assessments**. Wellington: NZCER.

Reid, N.A. (1978). 'Standards, Basics and National Assessment'. **Set: Research Information for Teachers**, No. 1, Item 13.

Assessment of Educational Achievement in Pakistan

- Policies, Implementation and Research -

Z. A. Ansari

Structure of Education and Examination

The educational system which is prevalent in Pakistan was introduced by the British in early 19th Century in selected areas of what lateron became British India, and still later divided into India, Pakistan and Bangladesh. In its present form a child passes through five years of primary education (subdivided now into a 3-year elementary phase and a two year post-elementary phase), followed by a three year middle and two year secondary phase. Education up to middle level is common for all the students, but at secondary level option of specialization in Arts, Science, Home Economics and Commerce is available. After Secondary school, a two years intermediate course is provided which is the basis for entry into specialized and professional education (e.g. Engineering/Medicine) or University education (BA or B.Sc. honours) or a further two year of college education leading to Bachelors degree in Arts, Science, Commerce or Home Economics.

With the introduction of new education system, the British also introduced an examination system on the model of London University of that time. The new system of Matriculation examination was first instituted with the establishment of the Universities of Calcutta, Madras and Bombay, in 1857. This was meant to be the first public examination of a student after ten years of education, and since it served as an entrance examination for the Universities, the Universities were made responsible for its conduct. Lateron it was realized that many students who pass Matriculation examination do not enter the Universities. Political and educational development led to the introduction of 'intermediate' stage - between the school and University education; and transfer of the responsibility of Matriculation and Intermediate examinations from the Universities to Boards of Secondary/Intermediate education. The second of these two developments was under process when Pakistan came into being in 1947. This division is now complete. Boards are responsible for all public examinations and policy matters regarding academic affairs up to grade 12, i.e. Intermediate. Universities confine themselves to graduate and post-graduate examinations.

Curriculum development takes place at national level. The present curriculum for grades 1 - 12 was developed in mid-seventies. A new round of curriculum revision is now underway.

The Boards

The Boards of Intermediate and Secondary Education (or BISE, as they are often called), are independent statutory bodies having broad powers in all academic matters including examinations. The boards have both elected and nominated members, the Chairman being one of the latter. The boards develop curricula, issue instructions about teaching, appoint head examiners and paper setters, and are responsible for conducting examinations.

At present there are thirteen Boards in Pakistan, located at Mirpur, Rawalpindi, Peshawar, Lahore, Sargodha, Bhawalpur, Multan, Hyderabad, Sukkar. There is a Federal Board for Islamabad and overseas schools, and an Armed Forces Board. All these boards cater for the requirements of both Secondary and Intermediate examinations. Karachi, however has two boards, one for Secondary School and the other for Intermediate examination. There are also technical boards responsible for the academic aspects of technical education.

The boards are coordinated by a committee consisting of Chairmen of all the boards. There is a great deal of coordination in the scheme of studies, curriculum contents, examination methods and even in time-tables for examinations. However each board retains a lot of independence in examination practices and frequencies of examinations.

A related institution is that of Directorate of education. The Directorate does not always correspond geographically with the boards. Its function is mainly administrative like recruitment, promotion and transfer of teachers. However some directorates are more active in academic field than others. While public examinations do not fall in the normal purview of the directorates, they do conduct scholarship examinations and some have instituted public examinations in their region

Public Examinations

Usually the first public examination is held after completion of ten years of schooling, and is called Secondary School Certificate Examination. This is a multiple-subject comprehensive examination, which is followed by Intermediate examination two years later. Some boards however administer the first public examination for SSC in two parts: part 1 at the end of grade 9 and part 11 at the end of grade 10. The final results of SSC depend on the combined marks of the two examinations. Similarly some boards hold the Intermediate examinations in two parts, held at the end of grade 11 and 12.

Some directorates hold public examinations at the end of 5th and 8th year. These correspond to voluntary scholarship examinations which are held at the same time. The purpose of these public examinations, for which no separate certificate is issued, has been described as practice and motivation for the students and checking the standards of the schools. There are wide variations in the contents, method and frequency of non-public examinations, depending upon the locality of school, qualification of teachers and system of supervision and monitoring. The most common pattern is to have at least two term examinations, both of them conducted by the class teachers during the course of the year, capped by an annual or promotion examination conducted by the class teacher but supervised, at least nominally, by the head. Some schools have a system of fortnightly or even weekly tests, but such schools are rather few.

Examination Reforms : Early Phase

Reforms in the system of evaluation of educational achievement stated in 1959 when the Commission of National Education (1959), popularly called the Sharit Commission considered this problem along with other educational problems and gave the first directions of change. The Commission observed that:

> "...The public examinations held at the end of classes X and XII are taken too seriously and absorb the entire attention of the teachers and the school authorities. Their shortcomings are notorious : They consist solely of written tests in which success can be achieved through mere memorization, and practically no effort is made to test the pupil's intelligence, and no credit is given for the work done during the two years course covering the examination. This completely destroys any incentive to study until the 'E-day' looms near, for it is too distant a goal to encourage substained effort over a long period".

The commission considered the possibility of doing away with the public examinations at the end of grade 10 and 12, but felt that they are also serving a useful purpose as "... they are a yardstick for comparing the merits of pupils from different schools and foster a competitive spirit among the examinees and between their schools".

The major recommendation of this commission was instituting internal evaluation among with public examinations. The Commission recommended that 25 percent marks should be allocated to internal evaluation and 75 percent to external public examination. The internal evaluation was to be based on fortnightly tests, along with ratings of pupil's conduct and character including cleanliness, truth-fulness, honesty and punctuality.

The Commission emphasized the need for research on examinations particularly in view of the fact that internal evaluation was being recommended, and a number of people had expressed the apprehension that this new system is likely to lead to malpractices. They recommended that the Secondary Boards should establish separate sections for carrying out monitoring and evaluation of the new system.

The system of internal evaluation was introduced at various levels but without much planning and preparation, and with almost no accompanying monitoring, evaluation and research. This resulted in widespread misuse and chaos. The Commission on Students Problems and Welfare (1966) had the following observations to make about this system:

> "···This system, ··· from what we have been uniformly told, completely failed to achieve its purpose. Evidence shows that this power has been widely abused both by teachers as well as educational institutions. Some institutions as a matter of general policy, we have been told, grant 23 or 24 marks out of a total of 25 to every student regardless of his performance in the class or the periodic tests. Few institutions maintain any record to class-work at all. It is only the students of reputable institutions who suffer under this system, as such institutions bestows upon them no special favour. If this power had been fairly exercised, it would have largely restored the respect and dignity that the teacher unhappily no longer enjoys. Its general abuse has resulted in making the position worse. It has led to unworthy students securing good marks at public examinations and unconscientious teachers claiming a record of unachieved performance for themselves. It is not surprising, therefore, that honest students and teachers alike should be unanimous in demanding that internal evaluation should either be abolished altogether or that at any rate it should not count towards the result of any public examination".

The early experiment having failed a number of educational reforms were abolished along with the system of internal evaluation.

Examination Reforms : Later Developments

In 1970, a period of great political turbulence in the country, the problems arising out of students unrest and its shadows on conduct of examination came into sharp focus. For the first time a National Committee was set up by the Ministry of Education and Scientific Research to examine specifically the examination system of the country, its current practices, problems and possible solutions. The report of the Committee (headed by Dr. M.A. Rashid, Chairman, East Pakistan Public Service Commission) was published in July, 1971.

The committee was given the task of reviewing the present system of examination in the country, comparing it with the systems of examination in other countries and making recommendations for the improvement of the system of examinations in Pakistan. However, very wisely, the committee did not confine itself to examination system and practices only. It took the view that the examination system was part of the educational process (which itself is operating in a particular socio-cultural milieu). The recommendations the committee made included provision of educational facilities, strengthening supervision in schools and improving the quality of supervision, setting up of provincial Bureaus of curriculum for keeping the curriculum and textbooks under constant revision, setting up a system of educational guidance and counselling, training of teachers in methods of evaluation and establishment of Research Cell for Test Construction and Evaluation at each university and board, and Provincial Bureaus of Testing and Evaluation for research as well as test construction. The committee emphasized the need for greater internal evaluation.

The committee recommended that:

1. All examinations at primary (grade 1-5) and middle (grade 6-8) should be internal, except for scholarship examinations.

2. In order to evaluate pupils on a continuous basis, Cumulative Record folder should be maintained for each student from class 1 onwards.

3. Periodic survey of the level of achievement of primary and middle schools should be undertaken by the District Education Authorities.

4. The first public examination (SSC) should be at the end of grade 10, and the next public examination (HSC) at the end of grade 12.

5. Both these examinations should be a combination of internal evaluation and public examination.

The committee also made a number of other recommendations on admissions, internal evaluation, supervision of examinations and marking of papers.

This report came out at the most turbulent time in the history of the country. It was therefore never formally implemented. However, the report still remains extremely important because most of the recommendations made in the report were accepted in one form or the other in subsequent educational policies and practices.

Most of the reforms proposed in the report of National Committee on Examinations found their way in the Education Policy 1972, which proposed a number of other changes in the examinations system:

(a) In place of single annual examination, a system of continuous evaluation of the progress, aptitude and problems of students by the class teacher will be introduced. Progression in primary classes will be automatic. Thereafter up to class IX progression will be based on a combination of periodical-cum-annual examinations and continuous, scientifically graded assessment of the students achievement, general behaviour and aptitude. For this purpose a cumulative record card of each student will be maintained by every school.

(b) The examination system after class 9 will be kept under continuous review and evaluation with a view to streamlining, revising or in any other way changing it.

(c) A system of giving special attention to students who do not show satisfactory progress in class work will be evolved aiming at no failures and no repetition of classes up to class IX.

(d) Children gifted with high intelligence and special aptitude will be identified from all over the country, and such gifted children will study in special types of schools which will be provided for such gifted and talented children.

(e) A massive shift of students from arts subjects to study of science and technical subjects depending on their aptitude for study of these subjects. The aptitudes will be assessed through the use of scientific tests and other instruments.

The guidelines given in 1972 Educational Policy regarding educational evaluation were strengthened in the National Education Policy (1979). The policy statement regarding evaluation and examination is as under:

> "it has constantly been observed that one of the fundamental reasons for the deterioration of educational standards is the perpetual continuation of an obsolete system of external examinations which encourages learning by rote and subjects the students to various stresses and strains. Efforts will, therefore, be made to improve the public system of examinations at the end of classes X and XII. Public examination at the end of class X will be gradually substituted by internal system of evaluation. The progress of students will be determined on the basis of periodic-cum-annual examinations and continuous assessment of the students' achievement, general behaviour and attitudes.
>
> The admission policies to higher education institutions and professional colleges will be substantially improved. Besides the marks of the public examinations, the marks obtained in internal evaluation along with the results of the aptitude and admission tests will also be given adequate weightage at the time of admission.

The improvement of examination system will entail a substantial change in the role of the Boards of Intermediate and Secondary Education from merely examining bodies to research-oriented professional organizations primarily concerned with the development and standardization of achievement, aptitude and admission tests".

Implementation of Reforms

In order to translate this policy into action, implementation plans were prepared by a number of committees, workshops and conferences. Experts were also invited from UK and USA who gave their opinions regarding methods of organizing the reforms (Deakin, 1974; Helmick, 1974). In order to bring out changes in examination system of the Universities, the University Grants Commission appointed a Committee under Moghni (1975).

During the early phase some of the changes were brought about too quickly and without adequate preparation. However, the pace has now been coordinated with the state of readiness of educational institutions. The major requirement for the implementation of these reforms were: (1) development of standardized tests, scales and inventories of students abilities, aptitude and achievement; (2) training of teachers in use of tests and evaluation instruments; (3) creating a strong research base for monitoring and evaluation of the programmes and for development of the evaluation instrument.

In 1974 a Test Development and Research Sector was created in the Curriculum Wing of the Federal Ministry of Education which launched an elaborate programme of test development with the cooperation of Institutes of Education and Research, Departments of Education, Boards of Intermediate and Secondary Education Departments of Psychology and Curriculum Development and Research Centres.

Since then a comprehensive programme of developing achievement tests has been undertaken. Item banks in the subjects of Economics, English, General Science, Physics, Chemistry, Biology and Mathematics have been developed for class IX-XII. Later on the items were reviewed and edited by experts and analyzed in terms of educational objectives. Model papers were prepared for paper setters.

The test item banks incorporate items of various kinds: multiple choice, completion, short answer and essay type. The tradition of complete reliance on essay-type test has been broken. For many years now the boards have been using short-answer and completion questions. This is going to be strengthened further by use of multiple choice questions, as the Inter Boards Committee has decided that from this year 25 percent marked will be allocated to a separately timed objective test.

A comprehensive programme of teacher training has been undertaken and a large number of workshops were held for training teachers in objective assessment techniques. The problem however is enormous. It involves not only developing knowledge and skills, but also change in attitude and motivation. The training programmes have however been successful to the extent that a base of trained manpower has been created. In order to further strengthen teacher training in evaluation, the new curriculum framed for CT and BEd programmes, now include papers on Educational Testing. Hence it is expected that fresh graduates from these institutions will not be in need of in-service training in testing and evaluation.

Cumulative Record Cards have been developed by a number of institutions and are being used, although the use is not as widespread as envisaged in the policy.

A number of research projects have been initiated by Universities, particularly Departments of Psychology and Education, by Boards of Intermediate and Secondary Education and by Curriculum Development and Research Centres and Institutes of Education and Research. In the area of achievement testing a major nation-wide research was carried out by the Primary Education Project of Ministry of Education, who developed achievement tests in selected subjects for students of class V, and administered them throughout the country. Most other studies have however remained confined to a limited area, usually a province or even a city, for obvious reasons of financing.

Apart from achievement tests, a large number of educational and psychological tests, including tests of aptitude, interest and personality have been developed. They include adaptations of Differential Aptitude Test, Thematic Apperception Test, Children's Apperception Test. MMPI, CPI, Otis Quick-Scoring Mental Ability Test, Columbia Mental Maturity Scale etc. Other tests developed by educationalists and psychologists locally include verbal intelligence tests, Urdu Adjective Checklist, Problem Checklist, Academic Self Concept Scale, Attitude Scales, and Questionnaire for Study Habits and Attitudes. Major researches have been carried out regarding identification of gifted children, rural urban differences, attitudes towards female education, use of aptitude tests for admission in various courses, comparison of the achievement level of Pakistani children with those of other countries, attitude of teachers towards new type of examinations, the frequency and nature of examination in schools and the attitudes of teachers and students towards such examinations.

Problems of Educational Research

Research on educational achievement is a relatively recent phenomenon. Although we find examples of excellent work on development of aptitude test and objective type testing in early sixties, research on a larger scale started in early seventies. This happened as a result of three factors: (i) The decision by Government agencies regarding development of tests of achievement and other related instruments and consequent provision of funds for this purpose; (ii) Establishment of specialist research-cum-teaching institutions like Institutes of Education and Research, and National Institute of Psychology; (iii) Greater use of educational achievement tests in schools and colleges.

The provision of funds to scholars and institutions has increased quite significantly during the last fifteen years, yet it has not reached the level which would ensure sustained and long term research, particularly at national level. The use of educational achievement tests along with other psychological and educational tests needs to be enhanced. Greater demand of tests would result not only increase in the number of tests but also in the quality of test. There is a need to have greater coordination between researchers within the country, because there is a lot of duplication in research work. Similarly, there is a need for greater coordination and cooperation between the countries of the region.

References

Deakin, J. Public Examinations in Pakistan: Report on a visit to Pakistan. Government of Pakistan, Ministry of Education. Islamabad, 1974

Examinations in Pakistan : Current Practices, Problems and Possible Solutions. Report of the National Committee on Examinations. Government of Pakistan, Ministry of Education and Scientific Research. Islamabad, 1971

Helmick, John, S. Report on Pakistan's Examination System. Government of Pakistan, Ministry of Education Islamabad, 1974

Moghni, S.M. Report of UGC study group on University Examinations. University Grants Commission. Islamabad, 1975

National Education Policy and Implementation Programme. Government of Pakistan, Ministry of Education. Islamabad, 1979

Report of the Commission on National Education. Government of Pakistan, Ministry of Education, 1959

Report of the Commission on Student Problems and Welfare. Government of Pakistan, Ministry of Education. Karachi, 1966

Assessment of Educational Achievement in Elementary and Secondary Education in the Philippines

M. M. Miguel

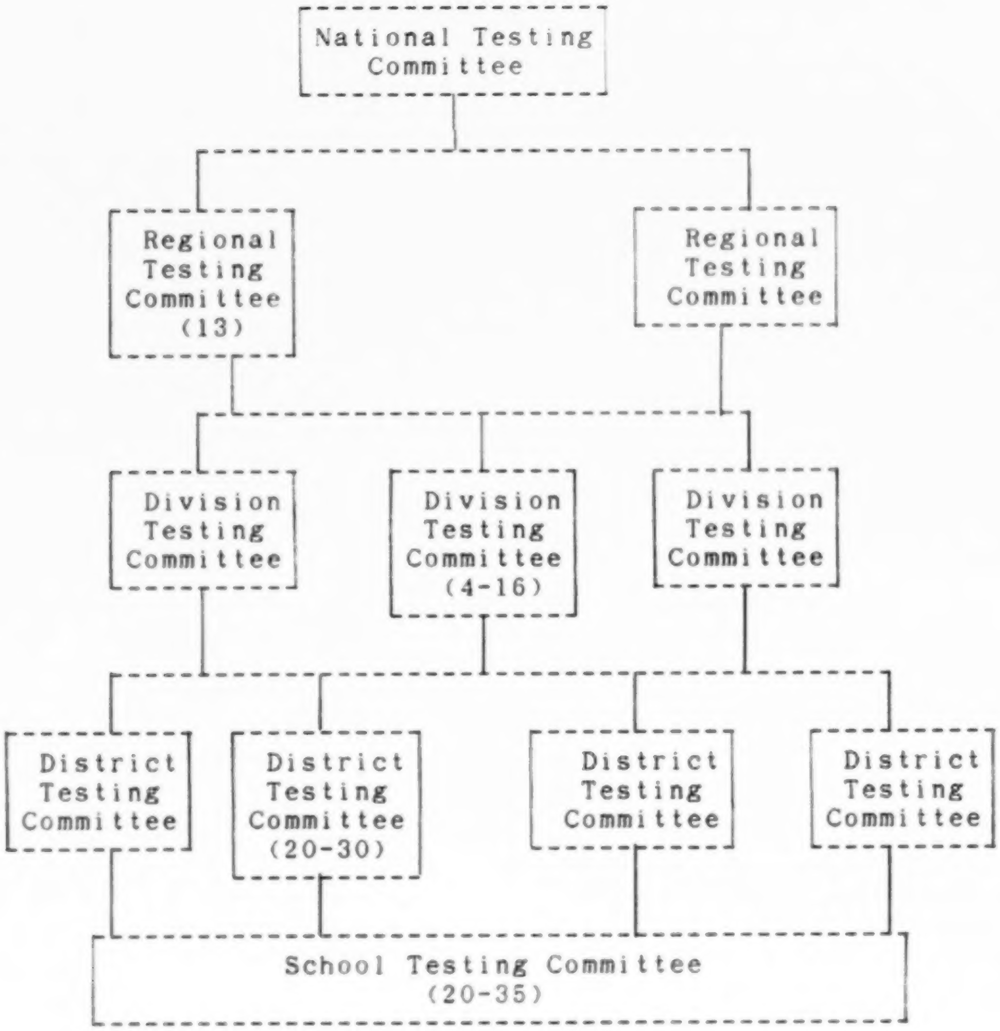
I. Purposes of Assessment of Educational Achievement

Assessment of student achievement is a regular component of the educational programme in the Philippines. Test results have long been used to communicate to students their achievement in school. Usually this is the grade the student is supposed to have attained and which serves as the basis for determining or deciding whether he passed or failed in the different subjects. In more recent times, the purposes of measurement and evaluation are being expanded and the emphasis is now being shifted to gathering of information about the students' readiness for instruction in a given subject area as well as his progress in the process of mastering what is to be learned. Data on which decisions are based relative to how effective the instructional programme in giving the student opportunity to learn maximally are also sought.

Mechanism for assessment in elementary schools

Passing a standardized national examination at the end of six years of schooling is not a requirement for graduation in the elementary level. The maintenance of a high standard of pupil achievement therefore becomes largely the responsibility of the teachers in the schools. Their capabilities in constructing or developing classroom tests as well as the analysis, interpretation and utilization of test results have been continually upgraded. There are still many teachers who do not fully grasp that evaluation is an important component of any goal-oriented instructional system, however.

Steps are now being undertaken to institutionalize evaluation or to ensure that it is recognized as a regular component of the total educational endeavour. In every school, a committee has been organized to plan and implement a programme for evaluating learning outcomes. Region-wide linkages among evaluation activities are being established through ad hoc testing committees organized on the school, district, division and regional levels as shown in the following figure.


National Testing Committee
Regional Testing Committee (13)
Regional Testing Committee
Division Testing Committee
Division Testing Committee (4-16)
Division Testing Committee
District Testing Committee
District Testing Committee (20-30)
District Testing Committee
District Testing Committee
School Testing Committee (20-35)

The Bureau of Elementary Education (BEE), a staff bureau of the Ministry of Education, Culture and Sports (MECS) initiates programmes on improving school assessment activities. For instance, a nation-wide training programme for school administrators was started in 1983 for purposes of orienting these personnel (regional and division research and guidance supervisors) so they can plan, implement and maintain their evaluation programme. They are expected, in turn, to train supervisors and principals in the different districts and divisions in the region. At present, the teachers learn about assessment of achievement in the School Learning Action Cell (LAC), the structure for institutionalizing the reorientation programme of teachers.

To coordinate efforts in the administration of diagnostic and achievement tests, MECS Order No. 33, s. 1983 titled, "Organizing A School Programme For Evaluation of Learning Outcomes" encourages the regions to do the following:

1) The regional office shall form a testing committee which initiate activities relative to gathering baseline data on pupil achievement levels a given year. Year-end achievement tests shall be given for at least 2 consecutive years to determine whether a target of at least 2% increase in achievement levels is achieved from the first (or baseline) year to the second.

2) The regional testing committee shall also be responsible for assessing the region's capabilities to achieve the region-based targeted increase in pupil achievement and how it fares in comparison to other regions in attaining this goal. The committee shall prepare a report of consolidated test data and findings together with their implications for policy-making, curriculum planning and improvement of instruction in the region and disseminate this throughout the region.

3) Schools, districts and divisions within the region shall be encouraged to prepare and administer their own tests and determine their own baseline data at the beginning of the school year which serve as starting point in instruction in the grade. Post-tests results at the end of this year will be the baseline for determining increase of pupil achievement at the end of the second year. Test results and their implications for curriculum planning shall be clarified in teachers meetings and conferences, and steps shall be taken for their implementation.

4) A nationwide evaluation of pupil achievement shall be jointly undertaken by the BEE and the National Evaluation and Testing Center to determine baseline data and increment of pupil achievement in terms of learning objectives and to make decisions about curriculum viability. The MECS then translates these into policies and programmes and implementing guidelines in improving the curriculum and the educational system as a whole.

Mechanism for assessment in secondary schools

The Bureau of Secondary Education (BSE) has organized an internal mechanism by which the quality of high school graduates is assessed, i.e., through an achievement testing programme that was initiated in 1981-82 and continuing to the present.

Assessment of student performance in the different subject areas of the curriculum is a necessary component of programme implementation. For this purpose, the Bureau of Secondary Education through the Curriculum Development Division has initiated a testing programme in 1982 at two points in the secondary education system, namely, the entry and the terminal points. Two kinds of tests were planned and prepared for administration of:

1) diagnostic tests for entering first year students; and

2) achievement tests for fourth year students.

The tests were administered through the joint efforts of the BSE and the MECS regional offices. Regional supervisors assisted the BSE representative in supervising the test administration in selected schools in each region. The test results were scored, analyzed and interpreted in the National Office of the BSE. Reports of the findings were sent back to the regions where these were discussed and their implications for policy formulation and curriculum improvements considered.

Likewise, high schools in the region, division and districts conduct assessments of achievement of students at the beginning of the school year for purposes of determining baselines in plotting student progress within the year. These diagnostic test results are needed in the targeting of expected increment in achievement levels of students, which is one of the more important factors considered in the performance appraisal system for teachers. This is true for both elementary and secondary levels.

II. The National Educational Testing Center (NETC)*

This is the agency which has been responsible for carrying out testing and evaluation functions for the MECS since 1975. Now in its tenth year of operations, the NETC is chiefly identified with the annual National College Entrance Examination (NCEE) for graduating high school students and high school graduates who desire to enroll in a degree course in college.

The NCEE programme has as its ultimate goals the following:

1) to check and/or minimize as much as possible any indiscriminate wastage of manpower and other resources which otherwise could be directed towards more profitable ventures, i.e., supplying the manpower needs of vocational, agricultural and technical fields;

2) to maintain a viable balance between the number of available white collar employees and skill workers in the manpower stock in this industry.

The NCEE is a general scholastic aptitude test which has been designed to provide an indication of the individual's potential to successfully do college work. It attempts to assess abilities and skills that develop generally and gradually over a long period of time from what one has learned both in and out of school. High school seniors and high school graduates from Philippines high schools and abroad should qualify in the NCEE if they wish to enroll in college courses requiring four years of study or more.

---

* Source: NETC Brochure. 1983

The NCEE is administered twice a year: the regular NCEE around September or October, in about 1,500 examination centres located all over the Philippines and also in ASEAN countries; and the Special NCEE on the last Sunday of June, administered in the cities of Manila, Cebu and Davao.

The first two NCEEs in 1973 and 1974 were conducted for the then Department of Education and Culture (DEC) (now Ministry of Education, Culture and Sports (MECS) by the Fund for Assistance to Private Education (FAPE). From 1975 and thereafter, the NETC assumed full responsibility for the development and administration of said examination.

The number of students who took the NCEE from 1973 to 1982 is summarized below. It will be noted that the number of examinees has consistently increased and has more than doubled since the inception of the programme.

Number and Percentage of NCEE Examinees Over the Years

Year	Regular NCEE		Special NCEE		Total	
	No.	%	No.	%	No.	%
1973	315,436	99.83	522	.17	315,958	100
1974	390,814	99.87	515	.13	391,329	100
1975	414,686	99.48	2,177	.52	416,863	100
1976	450,070	99.85	683	.15	450,753	100
1977	497,354	99.86	676	.14	498,030	100
1978	551,770 (Nullified)	99.89	589	.11	552,359	100
1979	569,799	99.37	3,634	.63	573,433	100
1980	663,718	99.03	6,470	.97	670,188	100
1981	700,436	99.43	4,002	.57	704,438	100
1982	746,098	100.00	-	-	746,098	100
Total	5,300,181	99.64	19,268	.36	5,319,449	100

Setting the NCEE Cut-Off Score

The Minister of Education, Culture and Sports, upon the recommendation of the NCEE Advisory Board, sets the NCEE cut-off score for a particular year.

The deliberation on the cut-off to be adopted considers the main thrusts of the Philippine educational system. These are Quality, Equity, Productivity, Filipinism, and Unity.

As regards education for Quality, it is admitted that at whatever percentile score in the NCEE the cut-off is set, some improvement in the quality of student input in college is achieved. The higher is the cut-off score adopted, therefore, the better relatively is the quality of students who enter college.

Along the Equity dimension, the principal consideration is whether those student-examinees who are financially disadvantaged but deserving of college education are in fact given the opportunity to pursue collegiate degree programmes. Provisions, therefore, are made to encourage this particular group of examinees to enroll in college.

As mentioned earlier, the NCEE programme is envisioned to redirect the flow of students from secondary to post-secondary institutions of learning so that with guidance, they would pursue relevant courses and eventually meet the manpower needs or requirements of the country's labour force.

On the thrust of Filipinism, the adoption of one national cut-off, it is believed, would prevent divisiveness and encourage a "concerned" attitude among the different regions of the country.

The following were the national cut-off scores adopted over the years:

School year 1974-75	25 percentile
School years 1975-76 to 1978-79	30 percentile
School years 1979-80 to 1981-82	35 percentile
School year 1982-83	40 percentile
School year 1983-84	45 percentile

(except for the cut-off for freshman enrolees in teacher-training courses which was set at 45 percentile in school years 1980-82, and at 50 percentile starting in 1982-83).

The Special NCEE

A Special NCEE is held every year, usually June following the regular NCEE. Those who are qualified to take the Special NCEE are:

1) Filipino students/scholars returning from abroad who were not able to take the regular NCEE. Filipino students currently enrolled abroad and will be graduation from high school in May or June, may be registered by

their parents or guardians.

2) high school graduates who took the NCEE but were not about to enroll in degree courses in college during the school year 1979-80.

3) applicants who took the Philippine Education Placement Test (PEPT), and on the basis of their test results, were accredited as having met the basic academic requirements for fourth year high school and therefore are eligible for admission to the next higher year level. They are to present their PEPT rating slips issued by the MECS.

4) those who registered for the regular NCEE but were not able to take the test due to illness on the day of the examination, as certified by a licensed physician or a reputable hospital.

Foreign students from the ASEAN countries desiring to enroll in Philippine colleges and universities also take the NCEE administered in their respective countries of origin. Students from non-ASEAN countries who similarly desire to take the Special NCEE are also served upon request.

The NETC also develops and administers the following:

1) Tests for screening applicants for the different scholarship programmes of the government, such as the State Scholarship Programme (SSP), National Integration Study Grant Programme (NISGP), Selected Ethnic Groups Educational Assistance Programme (SEGEAP), President Ferdinand E. Marcos Balik Kapatid Scholarship Programme (PM0697), TEIJIN, Pangulo Scholarship Programme (PSP), and Kabataang Barangay Scholarship Programme (KBSP).

2) The Professional Board Examination for Teachers (PBET);

3) The Philippine Educational Placement Test (PEPT) of the Accreditation and Equivalency Programme; (AEP)

4) The Philippine Specific Aptitude Test Battery (PSATB) for Guidance;

5) Assessment and other socio-psychological instruments like the Management Supervisory and Leadership Behavior Examination, Personality Tests and Cognitive Tests for school superintendents as well as for other school administrators; and

6) Specific tests and other measures to be used for specific programmes of non-government agencies, whose clientele are teachers and students both in the public and private schools.

The Accreditation and Equivalency Programme (AEP)

This is a project that is undertaken by the MECS through the NETC. It is geared towards the government's efforts to develop and utilize to the fullest human resources at all age levels.

It is a vehicle for giving drop-outs "a second chance."

The Rationale Behind the AEP

As early as 1972, there was a strong move to give increasing recognition to non-formal education primarily as a means of encouraging continuing study on the part of out-of-school youths whose circumstances in life prevented or restricted their continued participation in the formal school system. It was felt that education is an essential factor for achieving the social goal of equality and that as a result of education, the underprivileged gains an opportunity for vertical movement in society.

Objectives

In order to bring the ideas and concept underlying the AEP to a workable level, the programme was set to attain the following long range objectives: 1) to retrieve out-of-school youths and place them in the formal school system, if they so desire; 2) to accredit and validate knowledge and skills in academic areas gained through formal, non-formal and informal ways; 3) to accredit and validate work experience for purposes of job promotion, entry to job training, employment and/or self-fullfilment; 4) to develop and validate instruments to be used to measure skills in the various areas of practical arts as well as in the other occupational fields; 5) to develop a set of criteria for establishing equivalency.

Clientele Served

The AEP covers elementary and secondary levels of education. To be more specific, the following individuals are served by the programme:

1) Those who dropped out of school for a minimum of two years, provided that they are not placed at a grade/year level beyond that level in which they would have been classified had they gone through the regular formal schooling;

2) Those who dropped out of school and intermittently enrolled but did not finish any grade/year;

3) Those who have not gone to school at all;

4) Those who are presently in school, but over-aged for

their particular grade/year level and are recommended by the school authorities to take the PEPT. A deviation of not less than three years from the normal age for the particular grade/year level is considered over-aged.

5) The out-of-school youths and adults who previously took the PEPT but did not qualify for the next higher level.

Performance Profile of PEPT Examinees: From 1978 to 1982

Year	Total No. of Examinees	Retrogressed		Retained		Accelerated	
		N	%	N	%	N	%
1977-1978	9,091	No retrogression		2,426	26.69	6,665	73.31
1978-1979	22,727	13,974	61.49	2,700	11.88	6,053	26.63
1979-1980	21,121	11,224	50.74	2,778	12.56	8,119	36.70
1981-1982	26,050	9,695	37.22	4,064	15.60	12,291	47.18
Grand Total	79,989	34,893	43.62	11,968	14.96	33,128	41.42

Note: Test was not administered in 1980-81.

Testing Instruments

The objectives of the AEP are envisioned to be achieved with the use of two instruments: the Philippine Educational Placement Test (PEPT) and the Performance Test.

The PEPT is a battery of academic tests used to assess knowledge, skills and work experience which school drop-outs/leavers have acquired and which will be credited for academic equivalence on the basis of certain criteria.

The Performance Tests which are still in the final stage of refinement consists of test packages that span five areas of practical arts, i.e., agricultural arts, business and distributive arts, fishery arts, homemaking arts, and industrial arts.

The results of the PEPT and the Performance Tests will provide data for the proper channeling of school drop-outs whether for:

1) entry into the formal school system; or
2) training in some specific skills; or
3) job placement.

Registration and Test Administration

Registration for the PEPT is usually during September and October while administration of the test is normally held in December.

A school leaver/drop-out who is qualified to take the PEPT should register in the school division office nearest his/her residence and should present the following:

1) birth certificate;
2) last school card (Form 138); and
3) two ID pictures (size 1" by 1").

The school division offices will inform all applicants of the exact date, time and place of examination.

The NETC administers the PEPT through the regional offices.

The PEPT Certificate

The PEPT examinee gets the results of the test in the form of a certificate which stipulates the basic academic requirements he/she has met and the grade/year level he/she may now be admitted in.

The certificate which is signed by the Minister of Education, Culture and Sports and the Executive Director of the NETC takes the place of Form 138 (Report Card) issued by school heads.

Since its implementation from 1977-1978 to the present, the programme has already served 114,014 schools leavers broken down as follows: 9,091 in 1977-1978; 22,727 in 1978-1979; 22,121 in 1979-1980; 26,050 in 1981-1982 and around 34,025 in 1982-1983.

III. Major Researches/Surveys

1. Survey of the Outcomes of Elementary Education (SOUTELE), 1975

This was the Philippine first attempt at self-critical review of the performance of its elementary educational system. This national research project was designed to provide the decision-making machinery of the government as well as researchers, planners and implementers with the profiles of fourth and sixth graders. It also attempted to propose possible explanations of the relationship between scholastic achievement and school and non-school factors. The study provided the kinds of information necessary to make educational plans and researches in education better organized to contribute to a deeper understanding of educational problems.

The Project has for its main objectives the following:

1) To describe, measure and analyze the learning outcomes of sixth graders on a nationwide level in relation to the basic minimum learning package developed by educational experts for elementary education;

2) To identify factors in the learning environment both school and non-school, which could serve to explain the variations in scholastic achievement for various categories or groups of individuals;

3) To further identify from among these factors those which can be altered or modified through conscious planning and decision-making;

4) To recommend ways through which intervention can be affected, i.e., in terms of reallocation of resources or shift of emphasis in curriculum development, etc.;

5) To recommend areas requiring further research.

Project Design

It consisted of three major stages, namely, test development, test administration and data analysis and interpretation. Three sets of research instruments were developed for the Project. The first was a battery of achievement tests designed to measure the outcomes of elementary education in terms of the major curriculum areas: Language, Reading, Pagbasa, Wika, Mathematics, Science, Social Studies and Work Education. These tests were prepared with major objectives and specific learning outcomes of each area in mind. The second type of instrument used was a Student Attitude Inventory on values learned in Social Studies and Work Education. Thirdly, questionnaires were used to gather data on the social and economic characteristics of elementary school pupils, their teachers, their school heads and the facilities of their schools. The tests were administered to more than 20,000 sampled pupils across the country. The most significant findings are:

1) Elementary school graduates across the nation can answer only 50% of the test items correctly. The subject areas learned least are those on the 3 R's;

2) The elementary school pupils generally possess a positive attitude towards their teachers, school life and themselves;

3) School heads reported that faculties of their schools are at best "barely adequate";

4) The teaching and management staffs of the Philippine elementary education system are sufficiently trained;

5) On the average, teachers feel that students cannot cope with the learning requirements of their grade level;

6) On relationships:

   a) Pupil scores on the general mental ability test (Philippine Non-verbal Mental Ability Test) can be useful as an instrument for predicting the academic achievements of graduating elementary school students;

   b) Among intellectual skills learned in elementary schools, those learned in Science and Reading (English) are very similar or closely related to general mental ability;

   c) Socio-economic factors have a substantial impact upon levels of academic achievement;

   d) Attendance at pre-school and the type of elementary schools forms criteria for groups of pupils whose average achievement levels are significantly different from each other;

   e) There is a positive and substantial relationship between achievement scores and the socio-economic status of the pupil's families as measured by parental education and income;

   f) The teacher's evaluation of students is an important determinant of academic performance. As the teacher's evaluation of their students becomes more positive, the students' academic performance also increases above the levels of their own self-esteem.

It is worth mentioning here that the findings of Project SOUTELE formed the bases of far-reaching changes in elementary education in the Philippines which are now being implemented in the Ten-Year Development Plan for elementary education.

2. National Achievement Tests in PRODED (Program for Decentralized Educational Development)

With the implementation of the New Elementary School Curriculum (NESC) in PRODED starting in SY 1983-1984, a scheme to evaluate pupil performance was instituted in the different levels of the school system as directed in MECS Order No. 33, s. 1983. Project PREEM (Preparation of Elementary Education Evaluation Measures), a subcomponent of the PRODED-NESC project is responsible for evaluating pupil achievement on the national level.

Since the NESC is being implemented gradually, i.e. one grade per year, the scheme of evaluation is such that a grade implemented will be evaluated annually by Project PREEM in the first two years of implementation. The results of the achievement tests in Grade I, SY 1983-1984, administered by PREEM served as the baseline for determining the targeted 2 per cent increase in pupil performance vis-a-vis the Minimum Learning Competences in the NESC. With the second administration of Grade I National Achievement Tests at the end of SY 1984-1985, a comparative study of the performance of Grade I pupils is being undertaken. Consequently, the Grade II tests which are also being administered at the end of SY 1984-1985 for the first time shall be followed by a similar test in SY 1985-1986.

These achievement tests are given to a stratified sampling of pupils across the nation so the results can serve as a gauge for determining the overall performance level of pupils in the grade. The different regions are expected to establish their own baselines for measuring pupil progress down to the school level. If test results will be analyzed and interpreted in terms of levels of mastery of the objectives for each subject, the data will be valuable inputs in curriculum planning and development which include revision and modification if necessary.

Aims

The National Achievement Tests in PRODED are administered at the end of the school year for these purposes:

A. On Quality of Elementary Education:

1) To determine extent of mastery of minimum learning competencies in the curriculum;

2) To generate data which may be used as basis for improving the curriculum and instruction;

B. For Educational Planning and Policy Formulation:

3) To determine disparity in performance across regions, divisions and school types which could be basis for future investigations as to probable causes.

The Instruments

The National Achievement Tests in Grade I are in the learning areas: English, Wika (Language in Pilipino), Pagbasa (Reading in Pilipino), Mathematics and Sibika at Kultura (Civics and Culture).

The Sample

A nationwide sample of Grade I pupils was obtained through stratified random sampling. Seventy percent of the divisions in each of the thirteen regions represented by a total of 86 divisions were tested. Each division tested had a number of sample schools that approximate its relative size and which could be classified under the three kinds of school type (public central, public barrio and private) and the two kinds of school category (affected and unaffected).

Summary of Findings

1) Regional differences affected the pupils' achievement test scores, with the NCR surpassing the others in performance on all the tests.

2) It is evident that Central schools have a definite superiority over the barrio schools in terms of achievement in all the subjects in the new curriculum.

3) Although significant differences existed between affected and unaffected schools in the regional analysis, the direction of superiority was not consistent, so these differences tended to cancel each other out in the total population level.

4) As a whole, the Grade I pupils found the five tests moderately difficult, but within this average performance range, it was in Pagbasa that they did best and in Wika that they achieved the lowest mean score. Using 75% as criteria for mastery, it was found that all tests had less than half of their items and objectives mastered by at least 75% of the examinees.

Implications

1) The effects of locational differences were shown by the disparity in performance across regions, divisions and school types. Assuming that the curricular programme is unified and standardized, there may have been varied implementation procedures employed in the different locations. One highly probable variation would stem from the greater accessibility of teaching experts and instructional materials in areas like the regions and divisions closer to it, as well as in urban centres where central schools are located. A second possibility is the socio-economic background of the pupils which may intervene as a variable in their performance. These and other probable causes would have to be verified in future investigation.

2) The consistent results regarding school category effects on performance lead to the conclusion that a schools' rate of participation and retention do not matter in the pupils' achievement rates.

3) The pupils, in general, exhibit low mastery of the objectives specified in the tests for the different subject areas. There is a need to find out how these objectives were covered in the actual teaching-learning situation in order to explain the resulting mastery level. With regard to the mastery of skills in the cognitive domain, there is a need for the pupils to increase their achievement on all levels, including knowledge skills. The extent of mastery of the higher objectives shows that these skills can already be developed in the lowest rung of the educational ladder. Although at the start, the pupils do not achieve these to a superior extent, they can develop further in these skills even in their simpler forms of expressions.

3. National Diagnostic Tests in Secondary Schools

Purpose

A diagnostic test for entering first year high school students was administered by the Bureau of Elementary Education in SY 1982-83. The test was designed to reveal the strengths and weaknesses in learning possessed by these students and to acquire a performance profile of entering first year students.

Design

The tests were administered to 23,584 student drawn from the schools where the achievement tests for fourth year students of 1982 were given. The number represents 3% of the total school population of first year students. Subjects tested were Communication Arts (English), Science, Mathematics, Social Studies, YDT and Practical Arts. The instruments used are multiple choice tests on topics or content from Grades 4, 5, 6 and first year of the secondary school which was prepared to determine the stage of development and difficulties in learning of entering freshmen in different areas mentioned.

Findings

The major findings are:

1) National Capital Region (NCR) consistently got the highest regional mean in all subject areas except in Practical Arts.

2) The examinees performed best in Pilipino followed by Social Studies. However, they were deficient in Prac-

tical Arts, Mathematics and Science.

3) Across types of schools, students from private sectarian and city schools performed better than the students enrolled in other types of schools.

4) The skills in which the students were found to be strong were in knowledge of specific facts and terminologies and identification of classes. However, skills of the student in analysis, application and synthesis were found to be weak.

The study recommends the following:

1) Regions with low achievement levels should be given more assistance in terms of supervision, staff development and facilities.

2) School administrators and teachers should give more attention to the teaching of Practical Arts, Mathematics, Science and English.

3) Barangay and vocational high schools should be strengthened further in curriculum implementations.

4) The skills of analysis, identification and problem-solving should be emphasized in the teaching of Science.

5) For Communication Arts (Pilipino) reading and writing skills should be emphasized in classroom instruction.

6) Practical Arts should dwell more on realities and application of theories and principles learned.

7) YDT classes should train students in the process of decision making.

8) The Social Studies teacher should spend more time and effort in the development of the skills of comprehension and analysis.

9) Reading and Writing programme of high schools should be re-examined since these skills are needed in a science and technology-oriented world.

4. National Achievement Tests in Secondary Schools

Besides the analysis and interpretation of annual achievement tests, a comparative study of the different national achievement tests given to high school students was done by the BSE, other institutions and agencies, namely:

1) 1981 BSE achievement Tests;
2) 1982 BSE achievement Tests;
3) 1981 Diagnostic Tests of CEM (Center for Educational Measurement);
4) 1983 CEM Diagnostic Tests;
5) College Scholastic Aptitude Test (CSAT);
6) 1981 NCEE;
7) 1981 Philippine College Admission Test (UPCAT);
8) 1982 BSE Survey of the Public's Perceptions of Secon-Secondary Education.

From an analysis of the data from these different tests on student achievement, some of the important conclusions and recommendations were drawn:

1) The quality of high school graduates is dependent on the region from where they come from. Some of the regions are consistently below the mean, while some are consistently above the mean.

2) The quality of high school graduates varies depending upon the type of school they graduate from. Students from private sectarian schools and Provincial High Schools perform better than those from barangay high schools.

3) Variation in attainment at the high school level is related to attainment at the elementary level.

4) The high school graduate is proficient in Pilipino and Social Studies but deficient in Mathematics, Science and Practical Arts.

5) The scores of high school students in the aptitude tests (CSAT and NCEE) are declining, but their scores in the BSE Achievement Tests increased from 1981 to 1982.

6) Students tend to perform better in the lower level skills (knowledge level) than in the higher level skills (comprehension, analysis and synthesis).

7) In terms of aptitude, high school graduates have better reasoning abilities but are weak in verbal and reading comprehension.

8) High school graduates tend to be more articulate and involved in socio-political issues.

Recommendations

1) Some compensatory measures need to be initiated to enable the students from regions and school types that consistently perform below average to perform better.

2) Schools, divisions and regions should initiate plans designed to strengthen the competencies where students are weak and there should be close monitoring of progress.

3) Priority should be given to testing programmes at all levels of the educational system - school, division, region, and national levels.

4) Annual reports of schools, divisions and regions should focus first and foremost on the quantitative aspects of performance in school.

5) Studies/researches should be conducted to assess the quality of high school graduate in the Philippines with those of other countries.

Review of Research on the Short-Term and Long-Term Effects of the KEDI Educational System on Middle School

Joung-Kyu Whang

Introduction

The Korean Educational Development Institute (KEDI) has been conducting the Elementary-Middle School Development Project for the last ten years. One of the most prominent achievements of the Project has been that it has introduced a new educational system, i.e., the KEDI Educational System, under which the affective characteristics of the students as well as their academic achievement have improved. The introduction of the System has influenced education in Korea in many ways. Firstly, it has created the social environment for better schooling. Secondly, it has suggested policies for educational improvement. Thirdly, it has demonstrated the need for individualized instruction. Fourthly, it has urged the use of instructional television programmes. The achievement of the System has been proved through four small-scale tryouts and five large-scale applications. The management and the effect of the System have been favorably evaluated by the External Evaluation Committees at home and abroad.

In 1981, KEDI reported on the effect of the System in determining the academic achievement of the students.*1 According to the report, the students who studied under the KEDI System do better than other students showing a positive effect of the System on academic achievement. The report suggested a need for a further study to find out transfer effects of the System, that is, whether the cognitive and affective characteristics as well as the basic learning skills the students have developed under the System in their elementary school days can be transferred to middle schools. There had been a pilot study of the transfer effect in 1980.*2 The study compared the academic achievement of the first year middle school students who had studied under the System for four years in elementary school with that of those who had studied under the conventional system. However, this study used school test scores in measuring academic achievement. The interpretation of the results should therefore be made with

---

*1 KEDI, 1981, A Study on the Determinants of Academic Achievement of the Students under KEDI Instruction System.

*2 KEDI, 1980, A Study on the 5th Large-Scale Application of the KEDI Education System.

reservation, due to the possible lack of reliability and validity of the research instrument, i.e., the school tests.

Here, it is attempted to review a research on the short-term and long-term effects of the KEDI Educational System on middle school level in order to find out whether the System is effective in developing students' cognitive skills and effective characteristics.

Review of the Elementary-Middle School Development Project (E-M project)

The purpose of the E-M Project is to establish an educational discipline which will fit into the social milieu and respond to the demands of Republic of Korean society. The main task of the Project is to develop an educational system which will function in the educational context of Republic of Korea. The development of such kind of system involves a) setting educational objectives, curricula, and instruction at various school levels, b) designing an educational system which carries out the objectives, curricula, and instruction, and c) applying and revising the system. The KEDI Educational System is the product of such a procedure. The core of the KEDI Educational System is the KEDI Instructional Model. The Model has been designed to achieve intrinsic objectives describe the kind of experience the students will have when they are taught according to the Model. They prescribe the plan and the practice of the instruction and provide the criteria for evaluating the quality of the instruction.

The extrinsic objectives describe the long-term goals of the Model when it is implanted throughout the nation. They serve as the criteria for evaluating the Model itself, e.g., its practicability and effectiveness.

The prototype of the KEDI Instruction Model is shown in Figure 1. By 'prototype', it is meant that the Model can be adopted to individual subject considering the individualities of each subject.

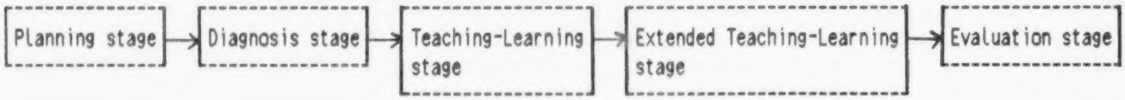


Figure 1. The prototype of the KEDI Instructional Model

The Model has five stages: Planning, diagnosis, teaching-learning, extended teaching-learning, and evaluation. The following describes the function of each stage.

Planning stage: to select a learning task and to plan instruction appropriate for the learning task.

Diagnosis stage: to diagnose the readiness of the students for the learning task and to provide appropriate preparation, if needed.

Teaching-Learning stage: to instruct according to the plan.

Extended Teaching-Learning stage: to evaluate the individual's achievement of the learning task and to compensate for under-achievement.

Evaluation stage: to evaluate the learning outcomes.

Each stage of the Model is accompanied by various teaching-learning materials.

The teaching-learning materials include teachers' manuals, student's workbooks, evaluation sheets, and broadcasting materials for instructional television or radio programmes.

The practice of the Model will be at its best when it is supported systematically by the schools concerned. KEDI provides a model for school management under the name of the KEDI School Management System. A school is viewed as an organization which plans, manages, and evaluates the elements of school administration, the most important of which is instruction.

Effects of the KEDI System on Development of Academic Achievement

The KEDI Educational System has been tested and revised through four small-scale tryouts and five large-scale applications with particular interest paid to the Instructional Model. The small-scale tryouts focussed on the instructional models for individual subjects derived from the prototype of the KEDI Instructional Model, and thereby evaluated the Model itself. The large-scale applications were meant to explore the generalizability of the findings of the small-scale tryouts and the effectiveness of the Model as a sub-system of the KEDI Educational System. The findings of the large-scale applications are reviewed below only with particular attention to the changes in academic achievement and the affective aspects of students.

The purpose of the large-scale applications were: a) to evaluate the effectiveness of the KEDI Educational System, b) to try out and revise the various teaching-learning materials, c) to diagnose the defects of the System and to design a counter measure for the defects, and d) to develop a teacher-training programme for nationwide implementation of the System. The large-scale applications covered 14 to 18 KEDI Cooperative Schools and 127 to 264 other schools. The duration of the applications was one to two semesters for all the school subjects.

The findings of the large-scale applications aré as follows:

a) The academic achievement

According to the result of the pre-test, which the entry academic achievement of the students in the first large-scale application, the experimental group and the reference group showed no significant difference in academic achievement.*1 For example, the mean scores of the 3rd grade were 64 for the experimental group and 65 for the reference group. For the 5th grade, the mean score for the experimental group was 57, being slightly lower than that of the reference group (M=61). The average scores of the 3rd and the 5th grades were 60 for the experimental group and 63 for the reference group. The result of the post-test conducted after the first application yielded a significant difference in achievement between the groups, the experimental group being the superior. The average score of the 3rd and the 5th grades was 62 for the experimental group (63 for the 3rd grade, 60 for the 5th grade), being 8 points higher than the 54 of the reference group (57 for the 3rd grade, 50 for the 5th grade). The difference between the groups became larger, the experimental group being the superior, over the five large-scale applications. That is, the longer the duration of the application, the larger the difference in academic achievement of the students between the groups. The changes in academic achievement during the application period are summarized below.*2

(1) There was improvement in the academic achievement of the reference group as well as of the experimental group. The case of the reference can be assumed to be due to a kind of audience effect or the expectancy effect, i.e., the reference group was aware of the fact that they were under study. Another possible explanation is that the rotation of the teachers might have contributed to the dissemination of the KEDI Instructional Model among schools. There is a possibility that teachers who were trained for the KEDI Model could have been transferred to other schools.

(2) There was a minor difference, although statistically significant, in the achievement of the lower grades, being in favor of the experimental group. It may be due to the fact that the duration of the application was shorter for the lower grades than that for the higher grades. If that is the case, it can be said that the longer the students learn under the KEDI System, the better they will achieve in school subjects.

---

*1 KEDI, 1976, A Report on the 1st Large-Scale Application of the KEDI Educational System.

*2 KEDI, Reports Nos. 36, 56, 71, 82 and 112.

b) The rate of mastery learning

In order to find out the effect of the KEDI System on mastery learning, the proportion of mastery students was compared between the experimental and the reference groups. Mastery students are defined as those who get more than 80 out of 100 in the academic achievement tests. The percentage of mastery students of the two groups is shown in Table 1.

Table 1: Percentage of Mastery Students during the Period of Large-Scale Application

Application period	Experimental Group (%)	Reference Group (%)
1st Application	27	16
2nd "	42	23
3rd "	45	20
4th "	59	32
5th "	62	12

According to the table, the proportion of mastery students in the experimental group has steadly increased, while that in the reference group shows irregularity. This can be interpreted as meaning that he KEDI System has reduced the number of non-mastery students, resulting in the overall improvement of academic achievement.

c) The reduction of regional disparity in academic achievement

According to the KEDI report on the academic achievement of elementary school students in Republic of Korea*1, students in the cities were superior (with a mean score of 66) to their counterparts in the rural areas (with a mean score of 55). The regional disparity was also reported in the result of the pretest of the first large-scale application*2, the city students being superior by 11 points. These reports point to the phenomenon of regional disparity in students' achievement as one of the problems that Korean education has to solve.

The report on the second large-scale application*3 showed an improvement of academic achievement of the students in the rural

---

*1 KEDI, 1974, A Report on the 2nd Small-Scale Tryout of the KEDI Educational System.

*2 KEDI, 1976, A Report on the 1st Large-Scale Application of the KEDI Educational System.

*3 KEDI, Reports Nos. 56, 71, 82 and 112.

areas up to 70, which was only 1 point lower than that of those in the cities. The application of the KEDI System has seen the increased achievement of students in the rural areas of the experimental group over the years, to the point where they are superior by 5-6 points to their counterparts of the reference group. This results in reducing the disparity of achievement between the students of the cities and rural areas to 5 points for the experimental group, while the reference group still retains a high degree of disparity at 8 points. These findings provide evidence that the KEDI System can contribute to the reduction of regional disparity in academic achievement.

d) The enhancement of higher mental abilities

One of the objectives of the KEDI Educational System is to provide the students with a learning experience that will enhance the higher mental abilities such as creativity, problem-solving, and evaluation. The results of the analysis are as follows:

(1) The mastery of the test items classified for the higher mental abilities

The experimental group was superior to the reference group in mastering the items classified for the higher mental abilities as well as for knowledge. The superiority was shown in all school grades and for all the school subjects. For example, the 5th and 6th grade students of the experimental group showed a high degree of mastery of the items classified for understanding and discovery which are closely related with higher mental abilities. In relation to the higher mental abilities required for school subjects, the experimental group was superior in those for Korean, Social Studies, Natural Science, Music, Fine Arts and Practical Skills. This can be interpreted to mean that the KEDI System helps students to improve their higher mental abilities.

(2) The development of the basic learning skills

The basic learning skills include the abilities of language use, arithmetic, and the use of materials which are required for school learning regardless of specific subjects. The comparison of the basic learning skills between the experimental and the reference groups is reported in the reports on the fourth and the fifth large-scale applications. According to the reports, the 4th grade students of the experimental group were superior in solving the learning tasks of their own school grade (M=74) to their counterparts (M=64) by 10 points. In doing the higher grade learning task, i.e., those for the 5th and the 6th grade levels, the experimental group was superior (M=68) to the reference group (M=57). The significance level of the difference is $p < 0.01$, the 5th grade students showed a similar tendency. This implies that the KEDI System provides the students with learning experience that will develop basic learning skills.

(3) Transfer of the acquired cognitive abilities to middle school level

The report on the 5th large-scale application studied the transfer effect of the learning abilities of the students who studied under the KEDI System in the elementary schools after they go to middle schools. The study compare the school test scores of the 1st year middle school students conducted at the end of school term between the experimental and the reference groups. The school test scores were transformed into T-scores by way of standardization. The experimental group was superior to the reference group by 3 points. This implies that the students of the experimental group accumulated learning abilities which were brought into play in attacking new learning tasks.

Effects of the KEDI System on Development of Affective Characteristics

The KEDI Educational System introduced student-centered learning and peer group learning in the place of teacher-centered teaching. It provides the students with such learning activities as survey, observation, discussion, and experiment. It is expected that through such activities students will develop a positive and active attitude toward learning and will acquire individualized study habits. The changes in the students' affective factors during the period of the small-scale tryouts and the large-scale applications of the KEDI Educational System are summarized below.

a) Self-study

One of the intrinsic objectives of the KEDI Instructional Models is to foster self-paced learning habits of the students. In order to achieve the objective, the students are provided with workbooks, the opportunities for discovery learning, experiment and peer group learning. These resulted in positive attitudes toward learning on the part of the students, e.g., there was a significant increase in the amount of self-study among the students of the experimental group. The reports on the fourth and fifth large-scale applications showed that there was a significant difference in the teachers' opinion of the students' groups, being in factor of the experimental group. The findings can be interpreted to mean that the KEDI System fosters the habit of self-study, which enables the students to manage their own learning.

b) Participation in learning

The KEDI Instructional Model provides the students with manageable learning tasks and various learning materials appropriate to the tasks. The effects of such efforts have been a) that the students' achievement have been improved, b) that the students' study habits and their attitudes toward learning have been enhanced, and c) that the students participate in learning with confidence and increased interest. The report on the fifth large-scale application has a word to say on the topic. After 72 classroom observations of the Social Science classes for the 3rd and the 5th grades in 12 schools (6 experimental, 6 reference), the report concluded that the experimental group participated in the classroom activities more actively, occupying 9% of the 40-minute instruction time, while the figure for the reference group was 3%. In question-and-answer sessions, the students of the experimental group were speaking 50% of the time, their fellows in the reference group only 40%. It shows that the KEDI Model encourages the students to participate in learning more actively and voluntarily.

c) The academic self-concept and the learning-related affect

The effect of the KEDI Model on the students' academic self-concept and the learning-related affect can be found in the reports on the fourth and fifth large-scale applications. The reports collected relevant data from the answers to a questionnaire distributed to the teachers concerned. According to the report on the fourth application, 74% of 505 teachers of the experimental group agreed that more than 40% of the students showed interest in learning, as against 41% of 246 teachers of the reference group.

In the fifth application, 92% of the teachers of the experimental group agreed that the students' confidence and interest in achieving the learning tasks had increased. The findings show that the KEDI Model has brought about an increase in the academic self-concept and learning-related affect of the students.

d) The cooperation among students

The KEDI Model provides the students with the activity of peer group learning, which is an attempt to introduce the cooperative management of learning among students themselves. According to the report on the fourth application, 94% of 493 teachers of the experimental group agreed that more than 50% of the students got used to peer group learning, in contrast to the 69% agreement of 213 teachers of the reference group. This is evident that cooperation among students has developed through the peer group learning among the students of the experimental group.

The Long-Term Effect of the KEDI System on Development of Academic Achievement and Affective Characteristics

In order to find out the long-term transfer effect of the KEDI System to the middle school level, a comparison was made in academic achievements and affective characteristics between middle school students who studied under the KEDI system as the experimental group and those who studied under the conventional system as the reference group. The specific questions raised are as follows:

(1) Do the cognitive abilities of study skills, which have been proved to be developed under the System in the elementary schools, continue to develop after the students go to middle schools?

(2) Do the affective characteristics of the students, which have been improved under the KEDI System in the elementary schools, continue to develop after the students go to middle schools?

The answers to the above questions will predict the transfer effect of the KEDI System and provide the rationale to revise the System, if needed.

The subjects of the study were 2,614 first grade middle school students (845 in the experimental group, 1,769 in the reference group). They were sampled by a stratified random sampling. 2 to 4 schools from each of the stratified regions (large cities, medium and small cities, and rural areas) have been randomly sampled. Among these schools, 5 girls' and 5 boys' schools have been sampled finally.

The variables investigated are academic achievement, attitude toward learning, and study habits. In the present study only the findings of academic achievement and attitude toward learning are summarized as below:

a) Difference in the academic achievement between the groups

Table 2 shows the difference in the achievement of the students between the experimental and the reference groups.

Table 2: Means and Standard Deviations of the Academic Achievement Test

		Korean	Social Studies	Math.	Natural Science	English	Overall
Experimental Group	Mean	70.71	60.29	69.73	65.03	68.68	66.89
	S.D.	13.15	12.57	15.55	14.29	17.54	14.62
Reference Group	Mean	65.52	47.43	60.07	54.68	59.95	57.53
	S.D.	15.50	15.70	19.86	18.15	22.36	17.86
Mean		+5.19	+12.86	+9.66	+10.35	+8.73	+9.36
F-test of Mean		*** 70.46	*** 433.17	*** 154.11	*** 211.62	*** 99.31	*** 261.10
S.D.		-2.35	-3.13	-4.31	-3.86	-4.82	-3.24
F-test of S.D.		** 1.38	** 1.56	** 1.62	** 1.61	** 1.62	** 1.49

** $p < .01$ *** $p < .001$

From the results shown in Table 2, it can be safely said that there is a tendency for the experimental group to be superior to the reference group in academic achievement and for the gap between the groups to become lager as the years go by the tendency suggests that the students of the experimental group have better acquired cognitive abilities such as understanding, application, analysis, and comprehension. In addition, it is highly possible that they have developed the higher mental abilities such as problem-solving, discovery, and creativity. If that is the case, it can be described as one of the transfer effects of the KEDI System.

Table 2 also shows that the standard deviations of the achievement scores of the experimental group are smaller than those of the reference group by 2 to 4. It indicates that the experimental group is more homogeneous than the reference group. The differences in standard deviations between the groups are rather large, and are in favor of the experimental group. In detail, the difference in S.D.'s for Mathematics, the learning tasks of which are related hierarchically, is rather large (4.3).

b) A Comparison of the Attitudes toward Learning between the Groups

Table 3 shows the means and the standard deviations of the scores of the attitude toward learning measured by a questionnaire.

Table 3: Means and Standard Deviations of the Questionnaire on Attitude toward Learning

		General	Korean	Social Studies	Math.	Natural Science	English	Total
Experimental Group	Mean	27.73	27.01	29.58	28.91	29.06	29.99	174.28
	S.D.	3.28	3.44	3.79	4.54	3.82	4.81	14.27
Reference Group	Mean	28.83	26.22	29.41	27.92	28.83	29.90	171.13
	S.D.	3.23	3.46	3.64	4.41	3.53	4.52	14.09
Mean difference (Exp.-Ref.)		.90	.79	.17	.99	.23	.09	3.15
F-value		** 8.93	** 8.31		** 6.27			** 7.14

** $p < .01$

According to Table 3, the experimental group shows more a positive attitude toward learning than the reference group. The attitude toward learning in general and the attitudes toward Korean and Mathematics and the total score of the attitudes are significantly higher (p .01) for the experimental group. The attitudes toward Social Studies, Natural Science and English are higher for the experimental group, but do not differ significantly. The significant difference of the attitudes between the groups is shown in Table 4.

Table 4: Attitude toward Selected School Subjects: A Comparison between the Experimental and Reference Groups

	General	Korean	Mathematics	Total
Experimental Group	29.73	27.01	28.91	174.28
Reference Group	28.83	26.22	27.92	171.13

The experimental group shows a more positive attitude. The difference is 0.9 for the attitude toward learning in general, 0.8 for the attitude toward Korean, and 1 for the attitude toward Mathematics. The difference is significant at the level of $p < .01$. The total score of the attitudes is the sum of the scores of attitude toward each subject and the attitude in general. The total score is also significantly higher for the experimental group. It suggests that the KEDI System encourages the effective characteristics of the students to be positive.

Suggestions for further studies

The KEDI Educational System has been developed and put into practice as one of the major tasks of the Elementary-Middle School Development Project. During the period of 1973 to 1980, four small-scale tryouts and five large-scale applications of the KEDI System have been undertaken. The present study has found that the effect of the KEDI System on academic abilities and learning-related affect can be transferred to middle school. Considering what the KEDI System has accomplished and what future contribution it could make, the following is suggested for further efforts to make the best of the System. Firstly, efforts should be made in designing teaching-learning materials according to the revised curricula. In 1981, when the KEDI System was at the stage of nation-wide implementation, the curricula for elementary and secondary education were revised, and the textbooks revised in accordance with the new curricula. KEDI took a leading role in revising the curricula and was in charge of writing the textbooks for the elementary and middle schools. The guidelines for revising the curricula and writing the textbooks reflected the valuable experience and outcomes obtained through the Elementary-Middle School Development Project. However, these can be regarded as a disciplinary side of effort on the improvement of education in Korea. In order for the disciplines to be activated, they should be embodied into a feasible shape. From this point of view, it is necessary to supplement the textbooks with subsidiary teaching-learning materials. The teaching-learning materials should include individualized learning materials with which students can diagnose their readiness for new tasks and reinforce what has been learnt. For teachers, there should be a handbook for writing summative tests which will guide the planning and the instruction as well as the evaluation. Secondly, a great deal of effort should be exerted to incorporate the affective characteristics of the students in schooling. As Bloom has summarized, the affective characteristics of the students explain 50% of achievement, the subject-related affect explains 10 to 17%, and the school-related affect explains 20%. Apart from the fact that the affective characteristics of affective factors is a key to the development of a human being as a whole person.

The National Institute of Educational Evaluation

Lim, Byung Lang

I. Background

The elapse of 40 years since the liberation in 1945 has brought an enormous development of education in Korea, owing to the democratic ideals of equal opportunities for education, continuous increase of people's interest in education, and the government's various educational policies to meet it.

One of the most striking trends in Korean education has been the quantitative expansion of school population. During the last four decades, the number of elementary school children has increased four times, and that of secondary school students 58 times. At the level of higher education, the number of students has increased more than one hundred and thirty times from eight thousand in 1945 to over one million in 1985, to reach the phase of universe in higher education.

However, educational development should be achieved not only in quantitative aspects but also in qualitative ones. It is true that the quality of education has sometimes been constrainted by the quantitative expansion in Korea. It also can't be denied that our educational policy at the national level was mainly carried out to cope with the problems caused by the expansion of education population.

In order to ensure the qualitative development of education, one of the tasks we should concentrate our effort on is the matter of educational evaluation, that is to say, to evaluate exactly the achievement of selected educational goals and the results. However, recently we are faced with some difficulties to surmount, such as the shortage of basic materials for the qualitatively control of education, various problems occurring in the process of educational evaluation, and the excessive human and financial requirement in evaluation management.

Accordingly, the switch-over to a more efficient and further reliable evaluation system has been a national task urgently required, and many educational problems are expected to be solved through the improvement of evaluation system, which will gain the confidence and trust of people, hoping for the equity and validity in educational evaluations such as National Examination for College Entrance, various qualification and selective examinations, students' scholastic achievement tests and aptitude tests.

For the efficient management of educational evaluation, an independent organization is necessary, which will be devoted to the enhancement of validity, reliability and objectivity in educational evaluation. To meet these ends, the National Institute of Educational Evaluation (NIEE) under the Ministry of Education was established in 1985.

II. Purposes and Functions

The ultimate purpose of the NIEE is to improve the quality of Korean education and to ensure the normal condition of schooling by implementing researches on educational evaluation and educational assessment projects at the national level. The specific objectives and functions are as follows:

1) To prepare for the increasing demands of educational evaluation

Various Examinations at the national level will be managed by the NIEE. In addition to the National Examination for College Entrance and the high school entrance examination, all kinds of psychological tests, and various examinations for the special purposes will be managed by the NIEE in the near future.

2) To improve the speciality of educational evaluation

The NIEE carries out a lot of studies, such as the research for the development of test items and the establishment of item bank, studies on the methods of educational evaluation for the improvement of educational evaluation in the elementary and secondary schools.

3) To take exclusive charge of educational evaluation services

Previously, educational evaluation services were managed by the Ministry of Education, local boards of education and some other institutions. From now on the NIEE will take charge of all kinds of educational evaluation services at the national level in order to utilize manpower, facilities and budget efficiently and effectively.

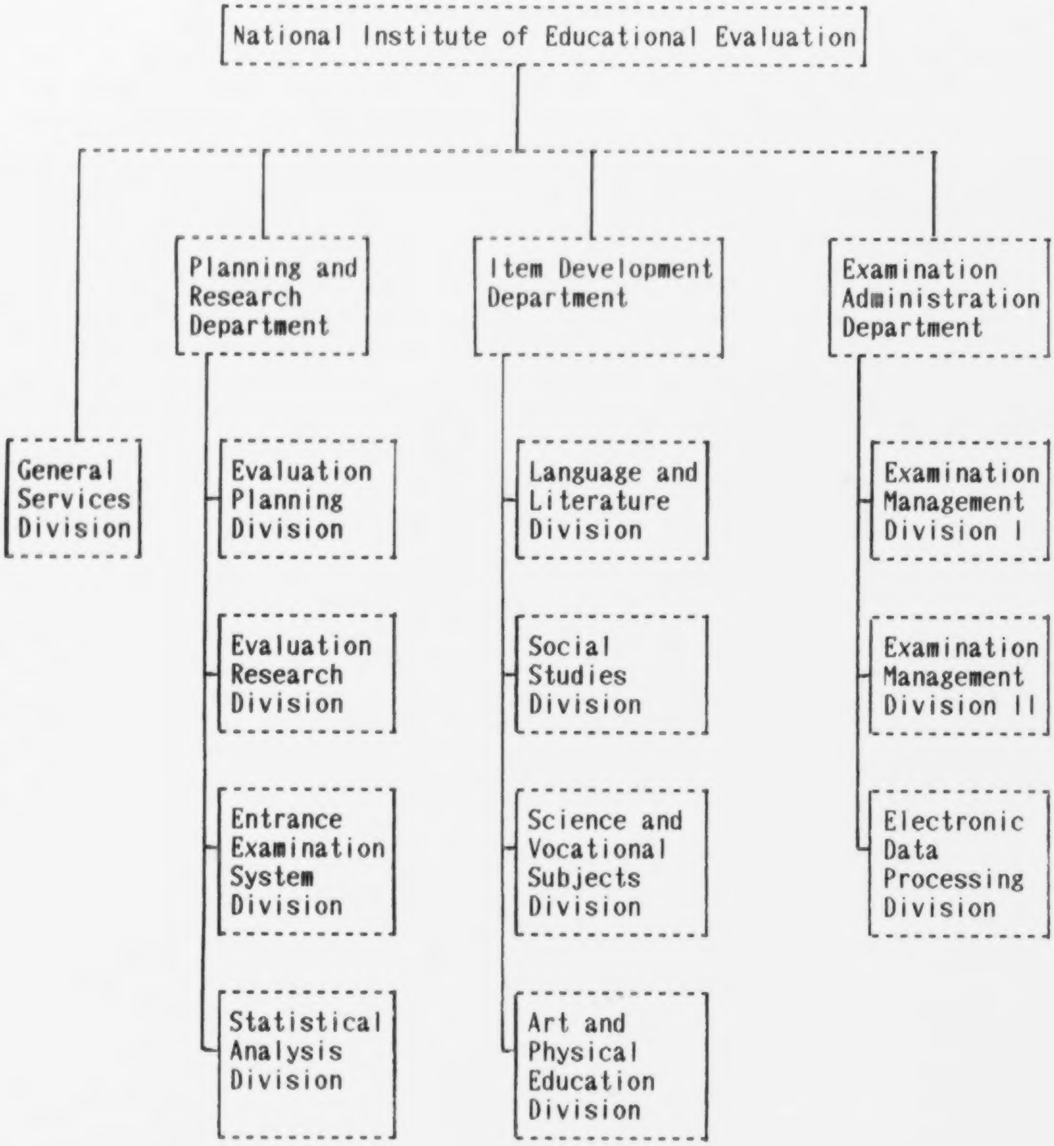
4) To provide basic educational information and materials for educational policy-making and supervision

The NIEE will store lots of study findings and data on educational evaluation to be used by educational policy-makers, supervisors and scholars.

5) To Support evaluation services in schools

The NIEE will help school teachers with the evaluation services such as the construction of test items and scoring through the operation of Item Bank System.

III. Organization


National Institute of Educational Evaluation
General Services Division
Planning and Research Department
Evaluation Planning Division
Evaluation Research Division
Entrance Examination System Division
Statistical Analysis Division
Item Development Department
Language and Literature Division
Social Studies Division
Science and Vocational Subjects Division
Art and Physical Education Division
Examination Administration Department
Examination Management Division I
Examination Management Division II
Electronic Data Processing Division

Assessment of Educational Achievement at Primary and Secondary Level in Sri Lanka

G.L.S. Nanayakkara

1. Introduction

Sri Lanka has a population of about 15 million composed of multi-racial and multi-religious groups. The majority of the people are Sinhalese who account for 74% of the total population. Sri Lanka Tamils constitute 12.6%, Muslims 7.1%, India Tamils 5.6% and others 0.7% of the population.

The educational system of the country is broadly divided into three sub-systems, namely;

(a) General Education;
(b) Vocational/Technical Education; and
(c) Higher Education.

General education is provided mainly within the formal school system. The school system in Sri Lanka is centrally managed. There are about 9,500 state schools and about 40 private schools.

The school education at all levels, except in private schools, is free. However, the management of education is primarily a state responsibility in the private schools as well.

Education in Teachers Colleges, Technical Colleges and Universities (except in private Medical College) is also free.

Medium of instruction in schools is Sinhala for Sinhalese children and Tamil for Tamil children. Children belonging to other races have the option of selecting the medium of instruction from Sinhala, Tamil or English.

The Ministry of Education is responsible for the design, implementation and maintenance of general education and teacher training programmes.

The Curriculum Development Centre (C.D.C.) of the Ministry of Education is responsible for all curriculum development activities. In-service training of teachers is done directly or indirectly by the C.D.C.

The Educational Publications Department of the Ministry of Education, with the help and guidance from University, C.D.C. and

Teachers College personnel, prepare school textbooks. These textbooks are distributed free to all pupils.

The entire system of public examinations is centrally controlled by the Department of Examinations of Sri Lanka, which comes within the purview of the Ministry of Education.

For administrative purpose the country is divided into 24 educational regions, manned by Regional Directors of Education. The actual implementation of the educational programmes is the responsibility of the Regional Director of Education, within his region.

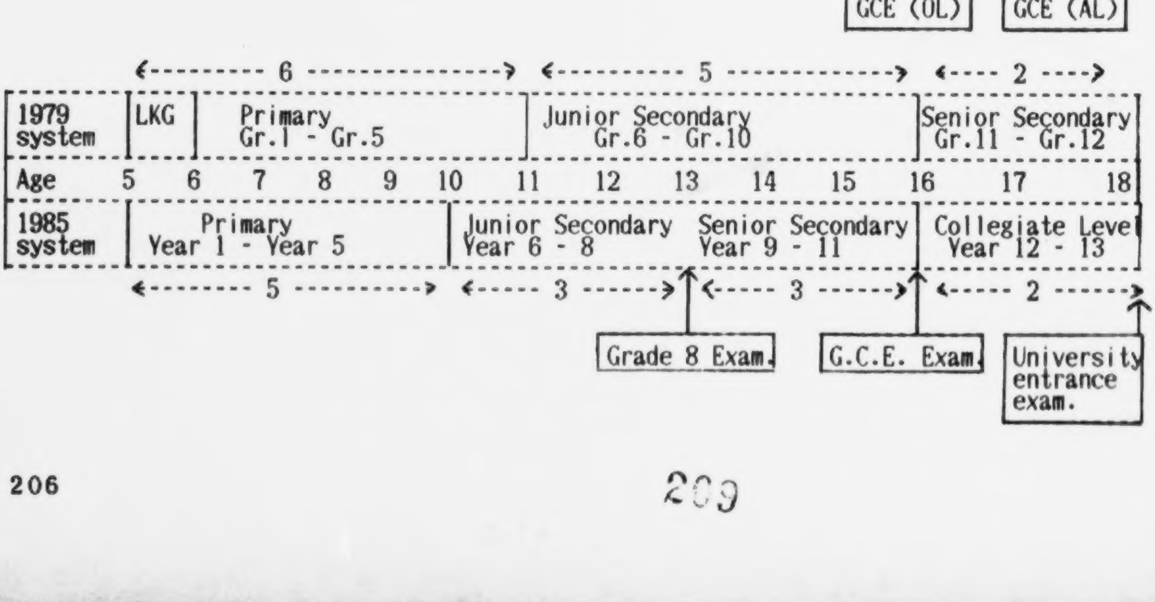
Each region is divided into several circuits. On the average each circuit consists of 30 to 40 schools. A circuit is in charge of a Circuit Education Officer (CEO) whose main function is the supervision of schools in the circuit. He also functions as the link between the schools and Regional Department of Education.

In addition there are specialist Circuit Education Officers who are in charge of supervision of various subject areas such as Science, Agriculture, Technical Subjects, Commerce, Home Science, Physical Education, Art, Music and Drawing. They are attached to the Regional Offices and are expected to keep regular contact with the relevant subject committees.

2. Structure of General Education

Several major reforms in the school structure and the school curriculum took place in 1972, 1977 and 1979. Another reform commenced in 1985 and is being gradually implemented in the school system.

The school structure which came into effect from 1979 and the new structure can be classified as follows:


GCE (OL)
GCE (AL)
6
5
2
1979 system
LKG
Primary Gr.1 - Gr.5
Junior Secondary Gr.6 - Gr.10
Senior Secondary Gr.11 - Gr.12
Age
5
6
7
8
9
10
11
12
13
14
15
16
17
18
1985 system
Primary Year 1 - Year 5
Junior Secondary Year 6 - 8
Senior Secondary Year 9 - 11
Collegiate Level Year 12 - 13
5
3
3
2
Grade 8 Exam.
G.C.E. Exam.
University entrance exam.

Note: This paper mainly deals with the systèm which came into effect in 1979, as it is still the major part of the system in force.

2.1 Classification of Schools

The majority of the Junior Secondary and Senior Secondary Schools have primary grades, and are mixed.

Schools according to Grades - 1982

Level	No of schools		
	Grade	Total	%
Primary	LKG - 5	4,007	42
Primary + Junior Secondary	LKG - 10	3,750	39
Primary + Junior Secondary + Senior Secondary	LKG - 12	1,415	15
Junior Secondary + Senior Secondary	6 - 12	372	4
Total		9,544	100

2.2 The School Curriculum

The school curriculum is designed and developed by the C.D.C. and is common to all schools in the country.

Primary Level:

The curriculum in the primary level is integrated in nature and teaching-learning take place mainly through activities. Here the main emphasis is on the provision of opportunities for the observation and appreciation of the physical environment, acquisition of basic skills in literacy and numeracy and understanding social life and elements of culture. The subsequent areas included in the curriculum are Religion, First Language, Mathematics, Environmental Studies, Aesthetic Studies, Creative Activities and English. Components of Health and Science education are incorporated in these subjects.

Junior Secondary Level:

This cycle consists of nine subjects, viz, Religion, First Language, English, Mathematics, Integrated Science, Social Studies, Health Science, Aesthetic Studies and Technical Subjects. Aesthetic and technical subjects consists of several optional subjects.

Senior Secondary Level:

This is divided into three streams, viz, Science, Commerce and Arts. At this level pupils have to follow a four subject curriculum. These four subjects are chosen from a large number of subjects.

Arts and Commerce - offer a choice of 50 subjects.

Science - Physical or Biological offer a choice of ten subjects in Science, Mathematics, Agriculture and Home Science.

3. Assessment of Pupil Achievement in Schools

3.1 Primary Level

Continuous evaluation of general performance of the pupil is the major procedure of assessing pupil achievement at the early stages of the primary cycle. In order to guide the teachers in continuous assessment, a special evaluation sheet has been supplied. Performances of the pupils in a class in a given subject can be recorded methodically in one of these sheets.

Achievement levels of a pupil in connection with different skills and concepts are recorded in the relevant column, in front of the pupils name. One sheet per subject can be used for the whole class.

Different sets of symbols are used to record the levels of achievement. The set of symbols used in Year I, Mathematics is given below as an example.

Pupil has achieved the expected ability level - ✓
pupil has not achieved the expected ability level - o
Ability level cannot be decided - ?

According to these symbols, a pupil who has not attained the expected ability level in a certain skill or concept at the first stage itself will be marked by the symbol - o.

If the child achieves the expected level at a latter stage it will be marked by the symbols - ✓.

A pupil who takes a longer period to attain the expected ability level, will be indicated by - ? ✓.

Model of an evaluation sheet is shown below.

Mathematics Evaluation Sheet - Year 1						
Skill or concept to be assessed						
Names of Pupils	Selects using one property	Selects using two properties	Observes patterns	Can identify several shapes	Can design a pattern using shapes	Measures using arbitrary units
1						
2						
3						
4						

Teachers could identify pupils who need remedial action, by constantly observing the performance levels recorded in the sheet.

In addition to continuous evaluation, mid-year and year-end tests are held to assess pupil performance, especially in the latter stage of the primary cycle. These are mainly paper and pencil tests. Question papers for these tests are usually prepared at circuit level or regional level.

Certain schools conduct monthly tests in addition to the above.

3.2 Junior Secondary Level

Various techniques are used to assess the pupil performance at this level and they vary among the subjects. However, the most common method of assessment is by paper and pencil tests. These test papers consist of objective type items and essay or structured essay questions. Every school is expected to conduct two major school tests, viz, the mid year test and the end of the year tests. Test papers for these tests are usually prepared at regional level.

Certain schools conduct monthly tests and term tests in addition to the above. The question papers for these tests are prepared by the teachers of the respective schools.

Assignments such as exercises given at the end of a lesson/ unit and tutorials are also used to a great extent to assess pupil performance in many subjects.

Practical examinations are held in certain subjects such as home science, needle work, aesthetic and technical subjects which cannot be assessed adequately by paper and pencil tests alone.

Children are encouraged to maintain resource book in certain subjects such as science and these are taken into consideration when the final assessment is made.

3.3 Senior Secondary Level

As in the Junior Secondary level the major procedure of assessment used is the paper and pencil tests. The question paper consists of objective type items, structured essay and essay questions.

At this level regional or circuit level tests are not widespread as in the Junior Secondary level. Certain schools prepare their own question papers while some schools combine together and prepare common question papers.

Regional level evaluation in mid year and year end school tests are being tried out in certain regions.

Assessment of pupil achievement through assignments such as exercises and tutorials play a major role at this level.

Assessment by oral questioning is another method used by many teachers, during classroom teaching.

At the school level practical work in science subjects is assessed by continuous evaluation method. In order to sit the public examination at the end of the course students should perform a pre-stated percentage of experiments and reach a minimum level.

Internal practical examinations are also conducted in some schools to assess the skills in practical work in science subjects.

3.4 Constraints on Teachers

There are many constraints on teachers which limit their ability to assess pupil achievement effectively. Some major constraints are listed below.

a) Large class sizes;

b) Heavy teaching loads of teachers;

c) Over loaded curriculum/examination syllabuses;

d) Pressure of the public examinations and the demand to cover the examination syllabuses within the specified period of time;

e) Inability to evaluate lab skills satisfactorily due to lack of laboratory facilities and equipment;

f) Teachers not trained adequately to design good evaluation instruments, to analyze marks and to make use of the results for remedial action and improvement of teaching;

g) Lack of resources such as stationery and duplicating facilities.

4. Public Examinations

4.1 Primary Level

An optional public examination is held once a year, for pupils completing the primary cycle, viz. Grade V scholarship examination. The main purposes of the examination are:

a) Selection of deserving students for the award of bursaries; and

b) Selection of students for placement in schools preferred by their parents.

This examination consists of two question papers set on First Language and Mathematics, duration of papers being 1 hour and 1.25 hours respectively.

This is a highly competitive examination and is taken seriously both by schools and as well as by parents. As a result mechanical drilling with model questions and private tuition are promoted. Many educationists are of the opinion that this examination distorts the teaching programme in upper primary grades and tends to defeat the educational objectives of the primary curriculum.

It has been observed that some of those children who are unsuccessful at this examination tend to feel that they are failures in life and this has a detrimental effect on their physical and emotional growth.

Another disadvantage pointed out regarding this examination is that the progress of the rural schools are slowed down as a result of admitting their best children, who excel at the scholarship examination to more popular urban schools.

Due to many reasons of the above nature, the Grade V scholarship examination has created a controversy among the public and many are of the opinion that this examination need to be modified.

4.2 Junior Secondary Level

At the end of the junior secondary cycle (Grade 10) pupils sit a public examination known as the General Certificate of Education (Ordinary Level), which is held once a year usually in December.

Main purposes of this examination are (a) selection and (b) certification. Admission to the senior secondary level as well as some of the vocational/technical programmes depends on the performance at the above examination. For the award of the certificate passes in six subjects are required, inclusive of First Language and Mathematics.

The method of testing at the G.C.E. (O/L) is mainly the written test type with practical tests in few subjects such as aesthetic, home science and needle work.

Two question papers are set in almost each subject and one out of them is usually an objective type test (fixed response or free response). Duration of the question papers varies with the subject. Structures of the question papers of few main subjects are as follows:

Subject	Paper I	Paper II
Science	40 fixed response objective items (4 responses) Time : 1 hour Weightage : 40%	10 structured questions 8 to be answered Time : 3 hours Weightage : 60%
Mathematics	30 free response objective type questions Time : 1 hour Weightage : 30%	12 structured questions 10 to be answered Time : 3 hours Weightage : 70%
Social Studies	40 fixed response objective type items Time : 1 hour Weightage : 50%	11 structured questions 6 to be answered (one question is compulsory) Time : 3 hours Weightage : 50%

Question papers are based on grades 9 and 10 content only.

4.3 Senior Secondary Level

At the end of the senior secondary cycle pupils sit the public examination known as the General Certificate of Education (Advanced Level), which is held once a year, usually in August.

The main purposes of this examination are also (a) selection and (b) certification.

Admission to universities and other higher educational/ technical institutes is a function of performance at the G.C.E. (A/L) examination.

There are two question papers almost in each subject and their formats vary according to the stream/subject. Their structure of the four science subjects, Physics, Chemistry, Botany and Zoology are as follows:

Paper I: 60 fixed response objective type items with 5 responses. Duration is 2 hours and weightage of marks is 50%.

Paper II: This paper consists of 4 structured essay questions (no choice) and 6 essay type questions out of which 4 have to be answered. Duration is 3 hours and weightage of marks is 50%.

Admission to universities is very competitive, the stiffest competition being for the science stream, notably the medical and engineering faculties.

School candidates can appear for the examination on two occasions only, after which they leave school and may appear as private candidates. Those who sit beyond the third attempt are not considered for university admission purposes.

A difference of a few marks in the aggregate of a candidate at this examination can make a vast difference in his social status when he steps out to the world of work.

The public interest in this examination cannot be underestimated. One index of this concern is well reflected in the large number of private tutors which have been sprouted in the recent past. Many a pupil go through the terminal years of his school career in an atmosphere of tension and insecurity largely due to uncertainty over what the outcome would be for them at this examination.

4.4 Setting of Question Papers

The question for public examination referred above are set by committees appointed by the Commissioner of Examinations.

Committees responsible for the setting of G.C.E. (A/L) question papers consist of university lecturers and a representative from the C.D.C., wherever possible.

The setters of the G.C.E. (O/L) question papers are chosen mainly from the C.D.C., Colleges of Education, Department of Educational Publications and from the education service.

In order to ensure that a wide range of content area is covered and sufficient questions are included to test various levels of achievement, a table of specifications is usually drawn at the stage of setting the question paper.

The question papers set by the said boards are subsequently moderated by another panel of moderators appointed by the Commissioner of Examinations, for each subject.

The services of the practising teachers are not obtained for setting and moderation work, as a precautionary measure to preserve the confidential nature of the examinations. Non-involvement of teachers for this type of work sometimes gives rise to various problems, such as questions which are either out of syllabus or of a more depth than expected appearing in question papers.

An item bank has been established in the Department of Examinations recently and a substantial number of items in three main subjects (viz. Science, Mathematics and Social Studies) at G.C.E. (ordinary level) standard have been developed, edited and pre-tested. However, the item bank is still unable to meet the demands of the public examinations.

4.5 Making of Answer Scripts

Large numbers of candidates sit the public examinations referred above. For example, about 500,000 appeared for G.C.E. (O/L) and about 150,000 appeared for G.C.E. (A/L) examinations in 1985. In order to reduce the time span involved in the release of results, a method of marking known as 'conference marking' is adopted.

Under this marking system, about ten to twelve marking examiners are assigned to a Chief Examiner. Marking examiners are teachers who are directly involved in teaching the relevant subject at the relevant classes.

Answer scripts are marked at several Evaluation Centres under the direct supervision of the Chief Examiner.

Detail schemes of assessment are prepared to achieve uniformity in marking, as far as possible. The usual practice is to prepare a scheme of assessment at the setting stage of the question paper. Various possible answers to each of the questions are discussed by the panel of setters and these are included in the scheme. These answers are then alloted marks.

After administering the examination, panel of setters and a selected team of senior marking examiners study the answers supplied by the candidates to various questions, in a selected sample of answer scripts.

The marking scheme is revised after this exercise, by suitably adjusting the allocation of marks and including various other possible answers, if necessary.

This procedure is repeated with a team of specially selected Chief Examiners, and after a prolonged discussion the marking scheme is finalized.

The results of the two main public examinations are released within a period of about three months and those candidates who are not satisfied with their results have the freedom to apply for re-scrutiny of answer scripts.

5. Researches and Surveys on Educational Achievement

Currently, there is no national co-ordinating body for educational research and dissemination. Researches and surveys in education are carried out at various levels mainly by:

a) Universities - by the members of the academic staff and by students leading for post- graduate degrees and diplomas in education;

b) Various departments of the Ministry of Education such as the evaluation and research unit, planning unit, C.D.C., department of examinations and department of educational publications;

c) Various educational associations and institutes such as Association for the Advancement of Science and Institute of Physics/Chemistry; and

d) Individuals or teams sponsored by various funding agencies.

However, systematic research studies in the field of educational achievement has not yet developed to a satisfactory level.

Brief notes on some researches and surveys/analysis which have been carried out on assessment of educational achievement by different categories of personnel, at different levels are mentioned below as examples.

5.1 Achievement of Primary Level Students in Reading and Mathematics - Grades 2 - 5

This research study was conducted by the Department of Educational Publications, commissioned and sponsored by the UNICEF and was carried out with the leadership of Dr. Mrs. T. Kariyawasam (Department of Educational Publications) and Dr. J. Wanasinghe (Education Faculty, University of Colombo).

Assessment of Educational Achievement

This is an attempt to assess and evaluate the achievement of primary school children from grades 2-5 in Reading and Mathematics. These are two areas to which great emphasis is given in the teaching programme of the primary grades.

Some of the major findings are as follows:

a) In Reading Tests, girls tend to perform better than the boys and the performance of urban children is superior than the rural sector.

b) Environmental advantages which some children enjoy throughout the school days is significant. Children from superior homes are likely to know the meaning of words long before children from homes without any cultural background.

c) The basic principles, which form the foundation for development of logical thinking are not comprehended by majority of children irrespective of the type of school.

d) Simple statistics and decimal fraction were beyond the scope of majority of the students.

5.2 Need Assessment Survey of Primary Schools in Mahaweli 'B' System

This survey was carried out by the Primary Committee of the Curriculum Development Centre of the Ministry of Education under the UNICEF assisted Mahaweli Development Programme, under the leadership of Mrs. C.M. Abayadeva, Director of Education (Primary).

One of the major objectives of the survey was to find out the achievement level of pupils in Grades 2 and 5 in First Language and Mathematics with a view to develop training programme for primary teachers in the area.

Major findings:

a) Majority of the children of Grade 2 are unable to formulate simple meaningful sentences and to spell and write correctly.

b) Students of Grade 2 have not been initiated into any type of creative writing.

c) Problems given in statements (instead of numerals) were found to be the most difficult for the pupils. This clearly reveals poor reading ability for similar problems given in numerals have been answered correctly, by the same child (Grade 2).

d) Familiarity with metric measurement is poor (Grade 5).

5.3 A Survey to Find Out the Performance of Grade V Scholarship Holders at G.C.E. (Ordinary Level) and G.C.E. (Advanced Level) Examinations

This survey was done by Mrs. Ranjanie Jawardena and Mrs. Hema Ranaweera, who were two senior teachers during the period of the survey.

Objectives:

a) To find out whether there is a correlation between Grade V scholarship examination results and G.C.E. (O/L) examination results.

b) To find out whether there is a correlation between Grade V scholarship examination results and the G.C.E. (A/L) examination results.

About one thousand scholarship holders in grade 6 in 1974 from three selected schools formed the sample.

They were traced upto Grade 10 or Grade 12 and their detail results of G.C.E. (O/L) and G.C.E. (A/L) were obtained.

In addition data was collected through a questionnaire.

Major findings:

a) There was a positive correlation between Grade V scholarship results and G.C.E. (O/L) results. The scholars had performed very well at the G.C.E. (O/L) examinations obtaining on average over 5 distinction level passes.

b) There was no correlation between Grade V scholarship examination results and the G.C.E. (A/L) examination results. There were cases where students who obtained 5 and 6 distinction level passes at the G.C.E. (O/L) examination had failed in all four subjects at the G.C.E. (A/L) examination.

5.4 Analysis of Performance at G.C.E. (O/L) - Science - 1984

This analysis was done by the controlling chief examiner of the G.C.E. (O/L) - Science, with the help of data supplied to him by the Chief Examiners who were in-charge of evaluation boards. Each evaluation board consisted of about twelve marking examiners. Data was collected by the marking examiners by selecting a random sample from the answer scripts which they assessed.

Choice percentages of responses, facility indices and discrimination co-efficients for each of the items in the Science I question paper which consisted of fixed response objective type questions were calculated.

Data from the Science II question paper was analyzed in order to determine features such as popularity of questions and performance level in each question.

Major findings:

a) Candidates have fared better in the content of Grade 10 than the content of Grade 9 (Question papers are set on content from Grades 9 and 10 only).

b) Performance in the items based on Physics content was remarkably low when compared with that in Chemistry and Biology items of Science I.

c) Very low percentage of candidates have chosen to answer the main question based on the Physics content in Science II. Science II paper consisted of a question each on Physics, Chemistry, Biology: three questions on Integrated Science and four other questions of different types. Out of those who attempted the question in Physics about 75% had been enable to score more than 4 marks out of 15 marks. According to these results a very low achievement level is apparent in the physics subject area.

5.5 Analysis of Multiple Choice Question Papers with Special Reference to Those Set at the G.C.E. (A/L) Examination

This analysis was done by a team headed by Prof. V.K. Samaranayake of Department of Statistics and Computer Science of University of Colombo.

It was undertaken as an exercise to help understand the value of as well as the problems associated with, multiple choice question examinations in measuring student ability or achievement.

The data available in the form of responses to 60 items of Physics, Chemistry, Botany and Zoology question papers of 1983 G.C.E. (A/L) examination were analyzed using classical analysis.

Major Findings:

a) Identification of items which should be rejected or revised before re-use. The identification was done on the following lines:

   - Rejected on response pattern.

- Rejected due to media-wise differences in response pattern.
- Rejected due to wrong translation.
- Rejected by discrimination index.
- Rejected by discrimination as measured by point biserial correlation.
- Rejected on one or more of the above counts.

b) Vital necessity for the introduction of a suitable guess correction, in marking MCQ question papers is apparent.

c) It is desirable to computer mark the answer scripts and simultaneously conduct a classical item analysis to obtain information such as test reliability, item difficulties etc.

d) If the test is on several distinct areas of subject matter, it is best that items in each of these areas are grouped as sub tests. Items within sub tests should be arranged in order of ascending difficulty.

5.6 A Statistical Analysis of the G.C.E. (A/L) Examination for Science Subjects

This analysis was done by Mr. K.L.G.S. Wijayasena, as a partial fulfillment of requirements for the Post Graduate Diploma in Statistics.

The main objective of the survey was to investigate the nature of the performance pattern of pupils at this examination. The investigation was carried out by taking into consideration the marks scored by all the pupils in the six subjects, Pure Mathematics, Applied Mathematics, Physics, Chemistry, Botany and Zoology, in respect of the G.C.E. (A/L) examination spread over the consecutive years 1982, 1983 and 1984.

Major findings:

a) Analysis of raw marks in the specified subjects clearly show that their distributions are not normal. The square root of these raw marks fit into a normal curve, where it is symmetrical about the mean. Thus a positively skewed beta distribution is valid rather than a normal distribution, when the scores of Physics are excluded.

b) The marks of the three populations, Sinhala, Tamil and English, which comprise the total population distribute themselves in an identical manner.

c) There exists a very high degree of correlation between subjects Pure Mathematics and Applied Mathematics, Physics and Chemistry, Botany and Zoology.

d) Consistency in question papers in the six specified subject areas was maintained over the three year period under review.

References

APEID report - Pupil Evaluation in Science

APEID report - National Strategies for Curriculum Development

Ministry of Education - Education in Sri Lanka

Ministry of Education - Education Proposals for reform

B. Premaratne - Examination Reforms in Sri Lanka

Sri Lanka National report - International Bureau of Education 38th International Conference on Education

Interim report - Committee to inquire into and report on public examinations at secondary school level in Ceylon

IUPAC and UNESCO Sponsored workshop report - Evaluation in Chemistry (August 1968)

Dr. (Mrs.) T. Kariyawasam and Dr. J. Wanasinghe - Achievement of Primary Level students of Sri Lanka in Reading and Mathematics (Grades 2-5)

Prof. V.K. Samaranayake
M.G.N.A.S. Fernando
A.P.S.R. Somasiri
G.N. Wickramanayake
Analysis of Multiple choice question papers with special reference to those set at the G.C.E. (A/L) examination

Mr. C.M. Abayadeva - Need assessment of primary schools in Mahaweli 'B' system

K.L.G.J. Wijayasena - A statistical analysis of the G.C.E. (A/L) examination for science subjects

Teachers Handbook - Year I

Controlling Chief Examiners Report - G.C.E. (O/L) Science 1984.

The Assessment of Educational Achievement in Thailand

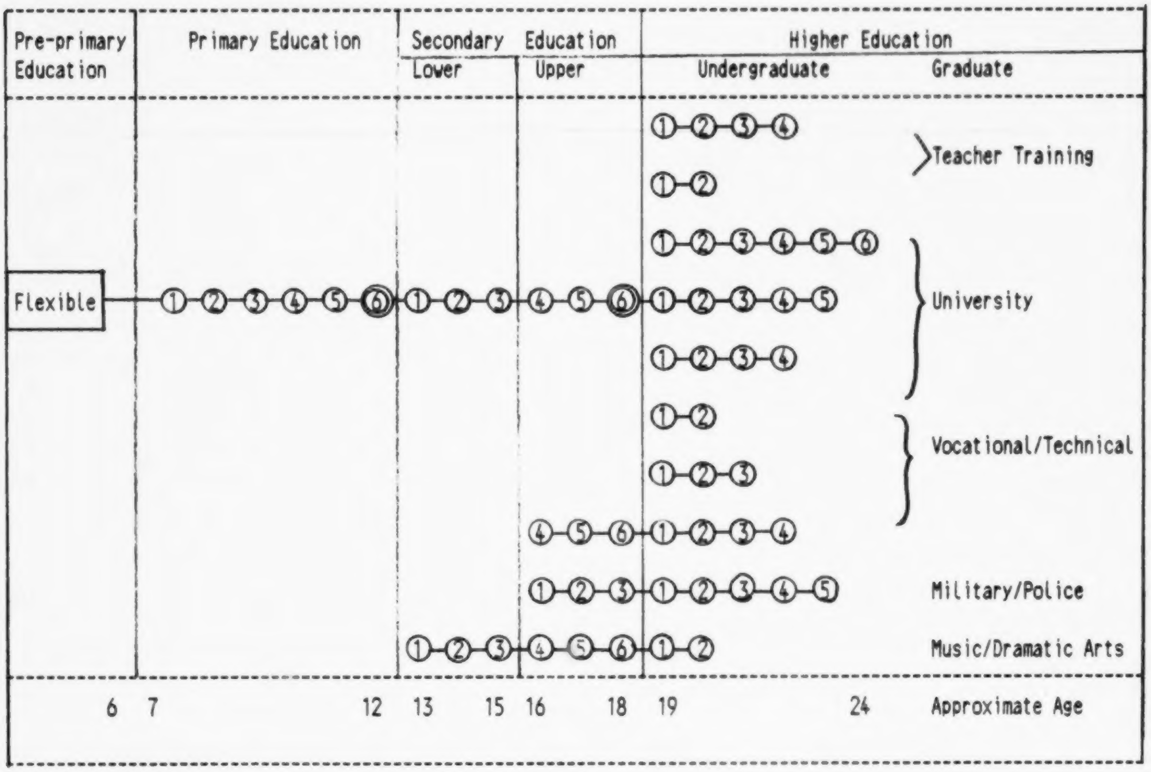
Winai Kasemsestha

A. Overview

1. School System

Based on the educational system, the formal school system is developed as illustrated in Chart I.

At the pre-primary level, the structure of school system is flexible, i.e., one year for pre-primary or two years for kindergarten. Approximate age of entry is 3-4 years old.


Pre-primary Education
Primary Education
Secondary Education
Lower
Upper
Higher Education
Undergraduate
Graduate
Flexible
Teacher Training
University
Vocational/Technical
Military/Police
Music/Dramatic Arts
6
7
12
13
15
16
18
19
24
Approximate Age

Remark: ◎ Terminal Grade

Chart I: STRUCTURE OF PRESENT SCHOOL SYSTEM

The primary education requires six years to complete. It is compulsory and free of charge. All children at the age of six are required, by law, to attend primary school.

Secondary educations is composed of two levels: lower secondary and upper secondary education, each of which requires three years to complete. At the upper secondary level, a student may choose either academic-oriented, or vocational-oriented programme, suitable to his aptitude and interest. A student may leave school system, if he so desires, after completing grade nine at the lower secondary level.

The school system at higher education is arranged in various forms. In general, upper secondary school graduates can pursue their higher education by taking entrance examination. Generally a university programme takes four years to complete, except Architectures (5 years) and Medical Science (6 years).

There are several specific higher education system modified to suit particular purposes of training, e.g., teacher training, technical education, music and drama, and military and police academies.

2. Curricula

Since the new educational system has been put into effect in 1987, the school curricula particularly at the primary and secondary school levels, have to be changed accordingly. The emphasis of the curriculum at each level is on self-completion, that is, students are provided with learning experiences such a way that they are well prepared for making a decent living if they are not pursuing to the next higher level of education.

(1) Primary Curriculum, 1978

This curriculum had been developed in accordance with the new school system of 1977. It aims at inculcating the learners basic knowledge and skills as stated in the National Education Scheme.

There are five groups of learning experiences; namely, Basic Skills (Language and Arithmetic), Life Experience (Social Studies, Environmental and Health Education), Character Education (Ethics, Arts, and Physical Education), Work experience (Pre-Vocation Orientation), and special group arranged for students at grade five and six (either English for Daily Life or Vocational Orientation).

(2) Lower Secondary Curriculum, 1978

The major objectives of lower secondary curriculum are to instill knowledge, skills, and appropriate attitudes towards earning a living. This curriculum is composed of five groups of learning experiences. They are : Language (Thai and foreign), Science and Mathematics, Social Studies, (e.g. Civics, Geography, History, Population Education), Character Education (Arts, Hygiene, Physical Education), and Work Experience.

In implementing the curriculum, the school employs a semester system. The time required is 40 week, 5 days per week for a total of three years to complete lower secondary education.

(3) Upper Secondary Curriculum, 1981

At the upper secondary school, students are provided with a relatively specific curriculum. Work-oriented learning experiences are given in parallel with those academic subjects. Subjects are composed of two groups, i.e., required and elective subjects.

Required subjects are those in academic groups:- Languages, Science, Social Studies, and Work Experience, which are arranged in accordance with major areas of occupations (Industrial Arts, Agriculture, Home Economics, Commerce, and Handicraft and Fine-Arts).

Elective subjects are offered so as to serve specific needs and interest of students.

(4) Other Curricula

There are also a variety of curricula designed for different programmes of education. For example, vocational and technical curricula, teacher training, and colleges and universities, each of which has its own requirements.

3. Administration of Education

Responsibility for the administration of education is shared among four ministries of national government: the Office of the Prime Minister, the Ministry of Education, the Ministry of Interior, and the Ministry of University Affairs. In general, the Office of the Prime Minister is responsible for overall planning, financial and staffing aspects of the whole educational system, the Ministry of Interior is responsible for primary education in municipal areas, the Ministry of Education's responsibilities fall upon most parts of the educational system ranging from pre-primary education to colleges at post-secondary level, while the Ministry of University Affairs is in charge of education at the tertiary level.

There are 14 departments within the Ministry of Education, each of which has its own specific functions. Half of these departments are either totally or partially assigned to operate or control certain types of schools, colleges, and institutes. Three such departments as Fine arts, Religious Affairs, and Cultural Affairs are responsible for developments in Fine Arts, Religions, and Cultures respectively. School curricula, textbook production, and tests and measurement of students achievement are under the Department of Curriculum and Instruction Development, while dealing with personnel matters are taken care of by the the Teacher Civil Service Commission. (see Chart II)

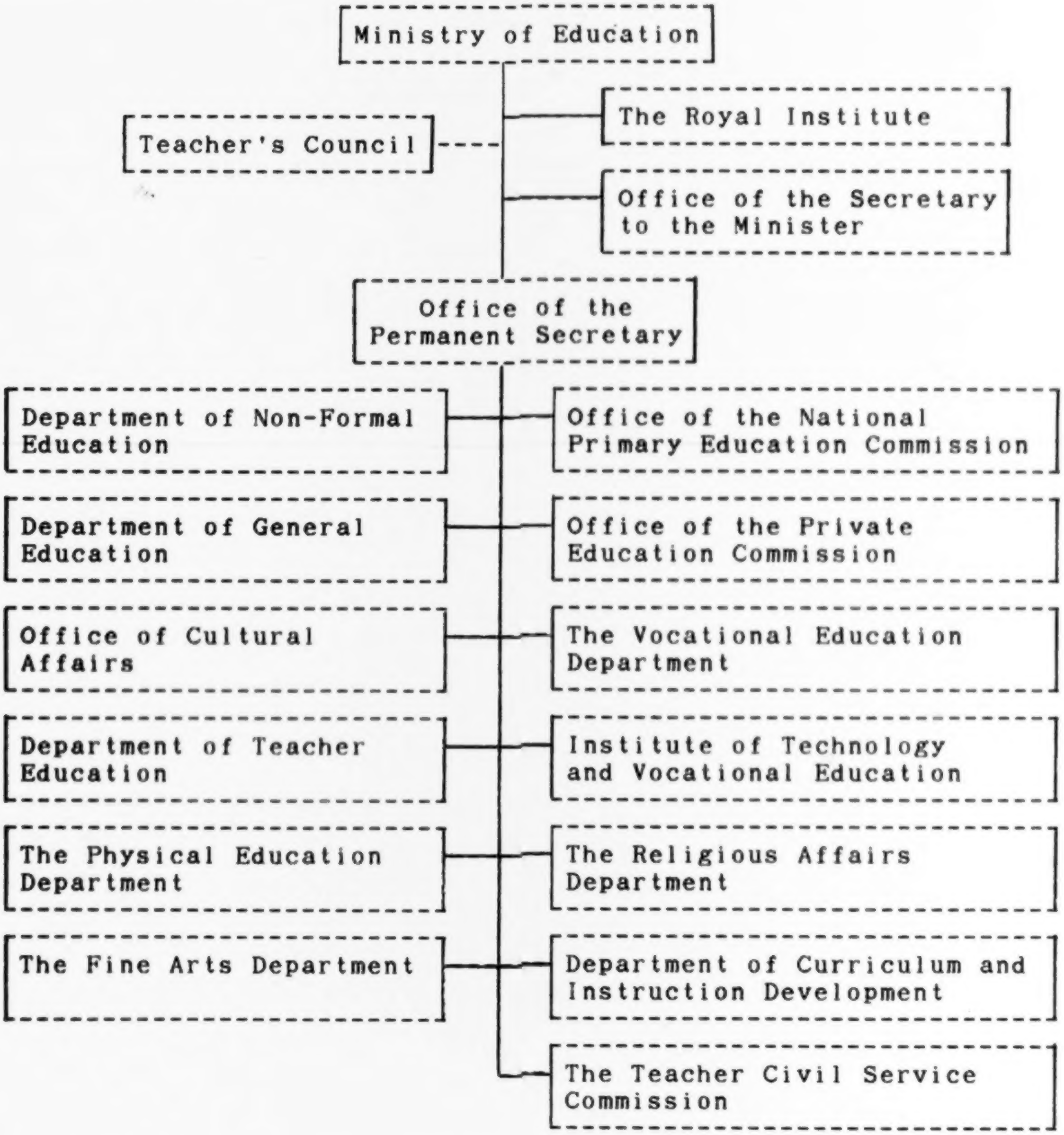

Ministry of Education
The Royal Institute
Teacher's Council
Office of the Secretary to the Minister
Office of the Permanent Secretary
Department of Non-Formal Education
Office of the National Primary Education Commission
Department of General Education
Office of the Private Education Commission
Office of Cultural Affairs
The Vocational Education Department
Department of Teacher Education
Institute of Technology and Vocational Education
The Physical Education Department
The Religious Affairs Department
The Fine Arts Department
Department of Curriculum and Instruction Development
The Teacher Civil Service Commission

Chart II: Organizational Arrangement of the Ministry of Education

Besides these 14 departments, the country is divided into 13 Educational Regions (Bangkok counted as one region). There are about 4 to 8 provinces in each Educational Region. Educational Regions except Bangkok, were set up for the purpose of improving the education in rural areas and to better adapt education to local needs as well as to geographical, occupational and cultural backgrounds.

B. Examination System in Primary and Secondary Schools

Following the principle of school accountability in the new curricula, evaluation of students' achievement for the purpose of classroom and course evaluation has been assigned to teachers with the approval of either principals or school cluster committees. The introducing of roles played by principals and school cluster committees is to establish some check and balance mechanism in the total system so that the high quality of learning outcomes is ensured. However, in actual practice principals and school cluster committees will intervene only in an informal manner when some serious problems that will affect the quality of learning are clearly evident.

To enhance the effectiveness of the instructional system, assessment of students' learning outcomes are conducted both during instructional sessions and at the end of the terms or semesters. In general within a term or a semester, teachers conducted three kinds of assessment. During a learning session or a block of lesson, the evaluation to see whether the students have mastered the specified objectives is conducted. If the students have not attained the accomplishment, remedial teaching is strongly encouraged. During the term or semester at least two structured summative evaluation sessions are provided to assess students' learning progress. Marks for these exams are added into the final course evaluation. At the end of the term or semester the final examination is given to evaluate over-all achievement of the course. (see Table I)

There is only a slight difference in practice between the course evaluation in primary level and secondary level. At the primary level, behavioral objectives are used and the passing of each behavioural objective is recorded. The percentage of passing a number of behavioural objectives is required to allow the students to move up to the next higher grade. But at the secondary level, learning objectives are used in both mid-term and final evaluation and the outcomes are recorded only from the final evaluation. However, secondary school teachers are strongly encouraged to use behavioral objectives arranging according to some hierarchical order from entering objectives to terminal ones, and teach and evaluate accordingly. Obligatory to remedial teaching are provided for those who do not obtain a satisfactory level in the evaluation.

Table I: Summary of Examination System in Primary and Secondary Education

Grade	In-school examinations	External examinations
1	Formative/Summative evaluation following the behavioral objectives (assumed mastery) up to grade 6	No
2		End of year exam with the approval of school cluster
3		No
4		End of year exam with the approval of school cluster
5		No
6		End of year exam conducted by district committees and National Assessment
7 - 9	Formative/Summative evaluation following course objective with remedial learning provided	National Assessment at Grade 9 and Standardized Achievement Tests provided for school uses
10 - 12	Formative/Summative evaluation following course objectives with remedial learning provided	National Assessment at Grade 12 and Standardized Achievement and Aptitude Tests provided for school uses

The organization of content areas are also differed between the two levels. In primary education, subjects are included in four learning experiences: the Basic Skills, Life Experiences, Character Education and Work Oriented Education. The summative and end of the year examinations are conducted according to these four groups of experiences. But, at the secondary level there are specific courses in each of the groups of subject matters. Hence, the summative evaluation are conducted for those specific courses in each semester.

A very large proportion of the test item used in these assessment are multiple choice question. Only a few are essay type such as, the writing skills and elaboratory skills. It is encouraged that the level of objectives should be high level thinking rather than memory skills. Nevertheless teachers are still using the low level of thinking skills and both affective and performance aspects are almost neglected. Attempts have been made to correct these drawbacks.

To ensure a high and equal educational quality for the whole Kingdom, external evaluation are added to the course evaluation. These external evaluation cover only some significant and more generic abilities and objectives. At the primary level the external examination considered as part of promotion and graduation are at the end of Grade 2, 4 and 6. The automatic promotion are used for Grade 1, 3 and 5. The assessment at the end of Grade 2 and 4 are conducted by schools cluster. For the end of Grade 6 the district committees headed by the District Educational Officer are responsible for the end of primary education assessment.

Another form of external evaluation is also conducted by the Ministry of Education aiming at the upgrading of the quality for the entire school system. The National Assessment are conducted either for the entire population or the sample during the last term or semester of Grade 6,, 9 and 12, the last grades for each level of education. The results of the National Assessment are not applicable to evaluation of the individual students for the purpose of passing or failing verdict. They are utilized for the overall evaluation and for planning of academic programmes at all levels, school, district, provincial and national.

In addition to the system presented above, there are also entrance examination for each of the connecting points of levels of education. Those who completed Grade 6 and would like to go on to the seventh have to take the entrance examination conducted by the school that they are applying for in case of the number of applicants is larger than the available seats. The same condition is also applied to those who finished Grade 9 going into 10 and those who finished Grade 12 going into tertiary level. The admission rates are about 45%, 98% and 22% for Grades 6 to 7, 9 to 10 and 12 to higher education respectively. For university education, there are two open universities that can enroll unlimited number and are not counted in the entrance examination and admission rate.

With these so structured assessment system, it is anticipated that the level of educational quality at all levels should be very satisfactory. But the results of the national assessment exercises so far still did not give a solid support in all aspects of learning outcomes. Therefore, the available resources in the Ministry will be thrown into the promotion of the educational quality programmes. They will be obligated for at least the next 5 years until the satisfactory level is achieved.

C. National Assessment

1. The System

The nationwide assessment of educational quality are conducted as part of either educational research programme or evaluation of curriculum or quality assessment. Currently three governmental agencies are actively involved in the programme. The National Education Commission (NEC) conducts researches that include students' achievement as main variables. Nationwide sampling is used to assure the generalization to the entire country. It is the purpose of the study to investigate variables that are contributing to the school learning outcomes so that some policy concerning educational resource allocation can be effectively made.

The Office of National Primary Education Commission (ONPEC) who is responsible for all governmental primary schools in the Ministry of Education conducts the national assessment programme to look at the quality of education. The results of the assessment will be used to promote the educational quality in every province and district. It is expected that the outcome from national assessment will be used in the achievement monitoring system to facilitate the educational progresses.

The third organization, the Department of Curriculum and Instruction Development (DCID) conducts all three programmes. The Educational Research and Planning Centre studies various practices that will lead to high quality of education. Field experimentation is utilized to cover the wide geographical areas and different settings. In addition, the Centre also conducts evaluation studies to assess the accomplishment of the new curricula so that some revision and improvement can be incorporated into the curricula for more successful implementation. The Office of Educational Assessment and Testing Services conducts annual national assessment from 1981. The more comprehensive one which covered the entire population of the 12th graders was in 1983. Sampling was used in 1984 exercise. The purpose of these assessment programmes is to report the level of educational quality to all agencies concerned in order to encourage some measures for the improvement of education.

Steps used in construction of items are similar among the three agencies. It begins with curriculum analysis to identify the expected key competencies. Students' overt behaviors are then specified for each of the competencies in every subject area. The items for the assessment exercise are corresponded to each of the sub-skills under each competency. Most of the items are multiple choice questions. However, some are short essay type to assess students' writing ability. It is desirable that the structural validity of these items must be clearly evident.

Examples of the results from the national assessment programmes are presented in Table II. It must be noted that the comparison among these results should take into consideration the fact that items used are different from one study to others. The purpose of the illustration is to give an example of what the outcomes are like.

Table II: Results from National Assessment Programmes from School Years 1979 - 80 to 1983 - 84 on Thai and Mathematics (number reported are percentage of Maximum scores)

Subject & grade	School Year				
	1980	1981	1982	1983	1984
3	DCID 54				
	NEC 52				
Thai 4	DCID 48		NEC 52		
6		ONPEC 35			ONPEC 49
9					DCID 45
12				DCID 52	DCID 47
3	DCID 45				
	NEC 51				
Math 4	DCID 50		NEC 52		
6		ONPEC 44			ONPEC 33
12				DCID 48	DCID 40

2. Results from Recent Studies

The Office of Educational Assessment and Testing Services conducted more extensive national assessment exercises in both 1983 and 1984 school years. The 1983 assessment covered both grades 6 and 12, and the 1984 one was for grades 9 and 12. In January 1986 the assessment shifted back to grades 6 and 12. Sampling was used for grades 6 and 9 but the entire population of the 12th graders were included in the 1983 and 1985 assessments. For the two studies in 1983 and 1984 the results could be summarily reported as follows.

In 1983 the sixth graders were administered five sets of tests covering Basic Skills (Math and Thai), Life Experiences, Character Education, Work-Oriented Education and Special Subject (English). The tests were included as part of the end of the primary education examination. All school districts were provided with a set of items. The evaluation boards at the district level would make their own decision to incorporate some items into the examination. The boards then were requested to send the answer sheets to the Office. There were two parts of the test, the cognitive and affective assessment. The results reported in Table III were only on the cognitive aspect. The affective part was mainly used in the developmental purpose.

Table III: Summary of Test Results (Cognitive Aspect) of Grade VI (1983 school year)

Subject	Full Score	Average Score	% of Full Score
Basic Skills			
- Mathematics	30	14.98	49.93
- Thai Language	50	29.28	58.56
Life Experience	40	19.95	49.58
Character Education	40	19.37	48.43
Work-Oriented Education	40	16.46	41.15
English	50	27.72	55.44

The results of the 1983 assessment indicated that the quality of education in the primary level that passed 50% mark on the instrument were in Thai and English. Life Experiences, Mathematics and Character Education were quite close to 50% mark, but the Work-Oriented Education was quite low, compared to the other subjects.

Table IV: Classification of Grade VI Test Results Using the Criteria of 40%, 60% and 80% of Full Score

Subject	% of the Students passing		
	40% criterion	60% criterion	80% criterion
Basic Skills			
- Mathematics	71.9.	26.49	3.71
- Thai Language	92.46	49.10	1.90
Life Experience	79.55	19.37	1.30
Character Education	76.51	16.34	0.77
Work-Oriented Education	54.18	2.76	-
English	79.13	41.43	8.57

For a more in-depth picture, Table IV illustrated the percentage of the students who passed 40%, 60% and 80% marks. It could be clearly seen that, again, the proportion of more than 40% of students who passed 60% mark were in Thai and English. It should be observed as well that the variation of students' ability was very diversed in English. The results of the assessment seemed to suggest that there was some problems in the Work-Oriented Education programme. Authorities concerned had to give more attention to every aspect of implementation.

In 1983 school year, the assessment of the 12th graders covered five basic subjects: Thai, Social Studies, Health Education, Mathematics and Science. The entire population was tested and the results presented in Table V.

Table V: Summary of Test Results (Cognitive Aspect) of Grade XII (1983 School Year)

Subject	Full Score	Average Score	% of Full Score
Thai Language	90	46.79	51.99
Social Studies	50	25.14	50.28
Health Education	20	10.47	52.35
Mathematics	60	28.79	47.98
Science	60	26.40	44.00

The areas that the students achieved higher than 50% level were Thai, Social Studies, and Health Education. The other two, Mathematics and Science were reported at 47.98% and 44.00% respectively. Assessing the achievement level against 60% mark, it was reported that the proportions of students who passed the mark were 17.7, 14.2, 23.9, 23.3 and 9.00 for Thai, Social Studies, Health Education, Mathematics and Science respectively. The results indicated that Science was the least successful programme among the five. Attempt should be made to upgrade the success of learning in Science.

It was also reported that disparity between the quality of education in Bangkok and rural areas was still very much in evidence. In general, students in Bangkok area performed much better than those in the rural areas.

Samples of the ninth and twelfth graders took part in the 1984 assessment. Some of the samples of the twelfth graders were intentionally included to motivate all administrators and students to put forward some measures to improve the quality of education. These samples were schools that performed at lower strata of the previous assessment in 1983.

The results of the assessment, 5 subjects for grade 9 and 6 subjects for grade 12 were reported in Table VI.

Table VI: Summary of Results of the National Assessments of Grade IX and Grade XII (1984 School Year)

Subject and Full Score*	Grade IX		Grade XII	
	Average Score	% of Full Score	Average Score	% of Full Score
Thai Language (60)	27.21	45.35	28.02	46.70
Social Studies (50)	22.45	44.90	21.70	43.40
Health Education (20)	11.72	58.61	12.72	63.60
Science (40)	18.26	45.65	-	-
Mathematics (40)	-	-	16.15	40.38
Work-Oriented Education (30)	15.28	50.93	16.18	53.93
English (50)	-	-	14.72	29.46

* Students in both grades sat for the same subjects but different contents (in accordance with the corresponding grades).

The 1984 assessment had included work-Oriented Education and English in addition to the 5 subjects conducted in 1983. Mathematics was provided for only those who took the intensive stream in Mathematics. The rest of the students would take English instead.

The results of both grades 9 and 12 showed that only Health Education and Work-Oriented Education had the averages higher than the 50% level. The lowest one was English with the average of 14.72 which was 29.48% of the total score. The rest were at 40% level. Considering the built-in bias stated earlier, it was concluded that the results of the assessment was very satisfactory. If more of better students were included in the samples, it was no doubt that the averages would be much higher.

The national report produced so far was only on the cognitive aspect of the assessment. The other part of the assessment, the affective aspect, was not yet to be reported to the public. However, many cautions had to be made before such national report could be disseminated. How to interpret the results and how to effectively utilize the results were two aspects that had to be made certain that teachers and administrators would really use then for developmental purposes.

Problems encountered during the implementation of the national assessment were as follows.

(1) There was a very large number of students to be assessed and every province and every school would like to have their students taken part in the programme. If a large group was covered, the monetary implication had to be met.

(2) The assessment had to cover every aspect of learning outcomes. So many subjects had to be included in the exercise of the assessment. It would take a lot of time in preparation of the exercise, the actual testing and the data processing. Therefore, it was decided that the exercise would be varied from year to year in term of competencies to be assessed and content areas to be used. It would be very interesting when the same item was used in the two to three years duration. If such situation was really occurred, the national progress could be clearly and validly illustrated.

(3) So far, there was little awareness and attempts to utilize the results of the assessment as a data base for school, district and provincial academic planning. These data produced by the assessment should be used to assess the school academic needs and these needs should be used, in turn, to formulate school academic programmes. Currently, efforts have been made to materialize these desirable impacts of the national assessment programme.

ANNEX

LIST OF PARTICIPANTS

Participants

Australia

Dr. Sid F. Bourke
Chief Research Officer
Australian Council for Educational Research (ACER)
P. O. Box 210, Hawthorn 3122
Victoria

China

Mr. Wang Yibing
Division Director
Policy Research Office
States Education Commission of China

China

Mr. Yang Dongliang
Officer
Policy Research Office
States Education Commission of China

India

Dr. P. N. Dave
Head and Dean Academic
Pre-school and Elementary Education Department
National Council of Educational Research and Training (NCERT)
Sri Aurobindo Marg
New Delhi - 110016

Indonesia

Dr. R. Ibrahim
Director
Examination Development Centre
Office of Educational and Cultural Research and Development
Tromol Pos 297 Kebayoran
Jalan Jenderal Sudirman - Senayan
Jakarta

Japan

Mr. Satoshi Kawanobe
Director
External Service Department
National Institute for Educational Research (NIER)
5-22 Shimomeguro 6 Chome, Meguro-ku
Tokyo

Malaysia

Mr. Mohd Rosli bin Mohd Ghazali
Head
Research and Data Unit
Examinations Syndicate
Ministry of Education
Jalan Duta
Kuala Lumpur

Nepal

Mr. Ramesh Jung Thapa
Additional Secretary
Ministry of Education
Kaiser Mahal, Kantipath
Kathmandu

New Zealand

Mr. Neil Alexander Reid
Chief Research Officer
New Zealand Council for Educational Research (NZCER)
Box 3237
Wellington

Pakistan

Dr. Zafar Afaq Ansari
Director
National Institute of Psychology
P. O. Box 1511
Islamabad

Philippines

Dr. Marcelina M. Miguel
Chief
Curriculum Development Division
Bureau of elementary Education
Ministry of Education, Culture and Sports
Palacio del Gobernador, Intramuros
Manila

Republic of Korea

Dr. Joung-Kyu Whang
Professor of Education
Department of Education
College of Education
Seoul National University
#56-1 Shinrim-dong, Kwanak-ku
Seoul

Republic of Korea

Mr. Lim Byung-Lang
Chief
Entrance Examination System Division
National Institute of Educational Evaluation
25-1 Samchung-Dong, Chongro-Gu
Seoul

List of Participants

Sri Lanka	Mr. G. L. S. Nanayakkara Teachers' College Lecturer Curriculum Development Centre 255, Bauddhaloka Mawatha Columbo 7
Thailand	Mr. Winai Kasemsestha Deputy Director-General Department of Curriculum and Instruction Development Ministry of Education Bangkok 10300

NIER Secretariat

Mr. Isao Suzuki	Director General
Mr. Takehide Yokoo	Deputy Director General
Mr. Satoshi Kawanobe	Director External Service Department
Mr. Toshio Kanaya	Chief Section for Educational Co-operation in Asia External Service Department
Mr. Miharu Kajita	Senior Researcher Ditto
Mr. Nobuya Higuchi	Researcher Ditto
Mr. Ryo Watanabe	Ditto
Mr. Akihito Kosakai	Associate Fellow attached to the Section for Educational Co-operation in Asia